

*Field effectiveness of stimulus equivalence for teaching reading skills to children with
Autism*

Lisa Giewercer, BSc.

Centre for Applied Disability Studies

Submitted in partial fulfillment
of the requirements for the degree of

Masters of Arts

Faculty of Social Sciences, Brock University

St. Catharines, Ontario

© 2010

Abstract

Stimulus equivalence involves teaching two conditional discriminations that share one stimulus in common and testing all possible conditional discriminations not taught (Saunders & Green, 1999). Despite considerable research in the laboratory, applied studies of stimulus equivalence have been limited (Vause, Martin, Marion, & Sakko, 2005). This study investigated the field-effectiveness of stimulus equivalence in teaching reading skills to children with Autism. Participants were four children with Autism receiving centre-based intensive behavioural intervention (IBI) treatment. Three of the participants, who already matched pictures to their dictated names, demonstrated six to eight more emergent performances after being taught only to match written words to the same names. One participant struggled with the demands of the study and his participation was discontinued. Results suggest that stimulus equivalence provided an effective and efficient teaching strategy for three of the four participants in this study.

Acknowledgments

I would like to take the opportunity to thank a number of people without whom this thesis would not have been possible. First and foremost I would like to thank the families and children who participated in my research. I would also like to thank the instructor therapists who worked directly with the children and were responsible for implementing all aspects of this research. Without their hard work and commitment, none of this would have been possible. I would also like to thank Julie Jones, one of the supervising therapists who oversaw children involved in the study for stepping forward and allowing this research to be completed. I would like to give a special thank you to the collaborators in this study. Namely, Denise Johnston and Heather Cushing Gordon for all their hard work in assisting in the implementation of this research, and Paul Szikszai and Melissa Lavallee for contributing to the development of the methodology and recommending participants from the teams they supervise. A special thank you also goes out to Dr. Nancy Freeman and Margaret Howard, from Surrey Place Centre, for their continuous support of this research and allowing me the flexibility to complete my degree. In addition, I would like to mention and thank my committee members, Dr. Tricia Vause and Dr. Barry Isaacs. Thank you both for all your feedback and support throughout the course of this project, your contributions were invaluable. I would also like to acknowledge Dr. Harry Mackay, who served as my external examiner, thank you for your feedback and suggestions. Your involvement in my defense, particularly your stories about working in Sidman's lab, made the experience invaluable.

My academic advisory, Dr. Rosemary Condillac, deserves special recognition for her invaluable support and contributions. Dr. Condillac pushed me and inspired me to get the most out of my education as possible. Without her, I would not be where I am today and I cannot thank her enough for her belief in me, her support, and her guidance. She is

an invaluable member of the Brock University Faculty and the field of Applied Behaviour Analysis as a whole.

Last but not least, I would like to thank my parents, in-laws, and siblings for all their support and patience with me in completing my degree. To my husband David, thank you for always believing in me and supporting me to achieve my dreams. I could not have done this without you.

Table of Contents

Abstract	ii
Acknowledgments	iii-iv
Introduction	1-2
Background	3-11
<i>What is Autism?</i>	3
<i>Outcomes of Interventions for Autism</i>	3-4
<i>Stimulus Equivalence</i>	4-6
<i>Reflexivity</i>	5
<i>Symmetry</i>	5
<i>Transitivity</i>	5-6
<i>Testing for Stimulus Equivalence</i>	6
<i>Challenges in Testing Equivalence with Auditory Stimuli</i>	6-7
<i>Stimulus Equivalence Research</i>	7-11
<i>Reading and equivalence</i>	7-8
<i>Experimental studies on Stimulus Equivalence</i>	8-9
<i>Applied Studies on Stimulus Equivalence</i>	9-12
Purpose	12
Research Questions	13
Method	13-41
<i>Participants</i>	13-15
<i>Participant 1</i>	14
<i>Participant 2</i>	14
<i>Participant 3</i>	14-15
<i>Participant 4</i>	15
<i>Setting and Staffing</i>	15-17

<i>Materials</i>	17-18
<i>Measures</i>	18-21
<i>Childhood Autism Rating Scale (CARS)</i>	18-19
<i>Scales of Independent Behavior Revised (SIB-R)</i>	19-20
<i>Baseline, Post Test, and Probe Session Data</i>	20
<i>Intervention Data</i>	20
<i>Inter-Observer Agreement</i>	20-21
<i>Treatment Integrity</i>	21
<i>Research Design</i>	22
<i>Procedure</i>	22-41
<i>General Procedure</i>	22-23
<i>Stimulus Equivalence Network Procedural Descriptions</i>	25-26
<i>Identity matching tasks</i>	25
<i>Production tasks</i>	25
<i>Matching dictated name to picture and dictated name to word</i>	25-26
<i>Remaining Match-To-Sample tasks</i>	26
<i>Absence of reinforcement</i>	26-27
<i>Staff Training</i>	27
<i>Procedure Steps</i>	27-29
<i>Word Selection</i>	29-31
<i>Baseline</i>	31-33
<i>Initial Test order</i>	32
<i>Revised Test Order</i>	32-33
<i>Intervention</i>	33-34

<i>Set up</i>	33
<i>Instructions</i>	33
<i>Stimulus Preference Testing and Reinforcement Procedures</i>	33-34
<i>Prompting Procedures</i>	34
<i>Error Correction Procedures</i>	34
<i>Mastery Criteria</i>	34
<i>Post Tests</i>	34-36
<i>Generalization</i>	36
<i>Follow Up</i>	36
<i>Error Analysis</i>	36-37
<i>Procedural Modifications for Individual Participants</i>	37-41
<i>Participant 1</i>	37-38
<i>Participant 2</i>	38-39
<i>Participant 3</i>	39
<i>Participant 4</i>	40-41
Results	41-97
<i>Summary</i>	41-46
<i>Participant 1</i>	46-61
<i>Reflexivity</i>	46-47
<i>Word Set One</i>	47-56
<i>Word Set Two</i>	56-58
<i>Word Set Three</i>	59-61
<i>Participant 2</i>	61-74
<i>Reflexivity</i>	61-62
<i>Word Set One</i>	62-70
<i>Word Set Two</i>	70-72

<i>Word Set Three</i>	72-74
<i>Participant 3</i>	74-86
<i>Reflexivity</i>	74-75
<i>Word Set One</i>	75-83
<i>Word Set Two</i>	83-85
<i>Word Set Three</i>	85-86
<i>Participant 4</i>	86-96
<i>Reflexivity</i>	86-87
<i>Word Set One</i>	88-96
<i>Rate of acquisition and efficiency</i>	96-97
Discussion	97-121
<i>Answers to research questions</i>	97-102
<i>Comparison to other studies</i>	102-104
<i>Extending other studies</i>	104-106
<i>Strengths</i>	106-109
<i>Limitations</i>	109-111
<i>Future Research</i>	111-113
<i>Conclusions</i>	113
References	114-121
Appendices	122-135
<i>Appendix A: Sample Data Sheets</i>	122-127
<i>Appendix B: Word Selection Results</i>	128-131
<i>Appendix C: Examples of Stimuli</i>	132-135

List of Tables and Figures

<i>Table 1</i>	<i>Participant description-CARS and SIB-R scores</i>	16
<i>Figure 1</i>	<i>Instructor therapist demographics information</i>	17
<i>Figure 2</i>	<i>The Stimulus Equivalence Network</i>	24
<i>Table 2</i>	<i>Words Selected For Each Participant</i>	31
<i>Figure 3</i>	<i>Summary graph for participant 1</i>	42
<i>Figure 4</i>	<i>Summary graph for participant 2</i>	43
<i>Figure 5</i>	<i>Summary graph for participant 3</i>	44
<i>Figure 6</i>	<i>Summary graph for participant 4</i>	45
<i>Figure 7</i>	<i>Generalization summary graph for participants 1-4</i>	46
<i>Figure 8</i>	<i>Follow-Up summary graph for participants 1-4</i>	47
<i>Table B2</i>	<i>Participant 1: Word Selection Results</i>	128
<i>Table 3</i>	<i>Participant 1: Reflexivity</i>	47
<i>Figure 9</i>	<i>Participant 1: Matching Written Word to Dictated Name</i>	48
<i>Figure 10</i>	<i>Participant 1: Matching Dictated Name to Picture</i>	49
<i>Figure 11</i>	<i>Participant 1: Matching Dictated name to Written Word</i>	50
<i>Figure 12</i>	<i>Participant 1: Oral naming of Word</i>	51
<i>Figure 13</i>	<i>Participant 1: Matching Picture to Written Word</i>	52
<i>Figure 14</i>	<i>Participant 1: Matching Written Word to Picture</i>	53
<i>Table B4</i>	<i>Participant 2: Word Selection Results</i>	129
<i>Table 4</i>	<i>Participant 2: Reflexivity</i>	62
<i>Figure 15</i>	<i>Participant 2: Matching Written Word to Dictated Name</i>	63
<i>Figure 16</i>	<i>Participant 2: Matching Dictated Name to Picture</i>	64
<i>Figure 17</i>	<i>Participant 2: Matching Dictated Name to Written Word</i>	65
<i>Figure 18</i>	<i>Participant 2: Oral Naming of Written Word</i>	66
<i>Figure 19</i>	<i>Participant 2: Matching Picture to Written Word</i>	67

<i>Figure 20</i>	<i>Participant 2: Matching Written Word to Picture</i>	68
<i>Table B6</i>	<i>Participant 3: Word Selection Results</i>	130
<i>Table 5</i>	<i>Participant 3: Reflexivity</i>	75
<i>Figure 21</i>	<i>Participant 3: Matching Written Word to Dictated Name</i>	76
<i>Figure 22</i>	<i>Participant 3: Matching Dictated Name to Picture</i>	77
<i>Figure 23</i>	<i>Participant 3: Matching Dictated Name to Written Word</i>	78
<i>Figure 24</i>	<i>Participant 3: Oral Naming of Written Word</i>	79
<i>Figure 25</i>	<i>Participant 3: Matching Picture to Written Word</i>	80
<i>Figure 26</i>	<i>Participant 3: Matching Written Word to Picture</i>	81
<i>Table B8</i>	<i>Participant 4: Word Selection Results</i>	131
<i>Table 6</i>	<i>Participant 4: Reflexivity</i>	87
<i>Figure 27</i>	<i>Participant 4: Matching Written Word to Dictated Name</i>	88
<i>Figure 28</i>	<i>Participant 4: Matching Dictated Name to Picture</i>	89
<i>Figure 29</i>	<i>Participant 4: Matching Dictated Name to Written Word</i>	90
<i>Figure 30</i>	<i>Participant 4: Oral Naming of Written Word</i>	91
<i>Figure 31</i>	<i>Participant 4: Matching Picture to Written Word</i>	92
<i>Figure 32</i>	<i>Participant 4: Matching Written Word to Picture</i>	93

Introduction

It has been argued by many that reading is essential to both success in school and a functional life in society (e.g. Rinaldi, Sells, & McLaughlin, 1997). Learning to read not only increases learning opportunities but also can increase communication opportunities (Nation & Norbury, 2005). For the majority of children, learning to read is a relatively easy process that occurs by mid-childhood (Nation & Norbury, 2005). However, for some children, especially those with disabilities, reading requires the integration of a set of complex skills and is anything but easy.

While very few studies on the reading skills of children with Autism have been conducted, it is known that many children with Autism are unable to read (Nation & Norbury, 2005). Difficulty with oral language has been linked to risk for literacy failure (Catts & Kamhi, 2005) and it is therefore not surprising that children with Autism struggle with these skills (Nation et al., 2006). For those children with Autism that do learn to read, failure to acquire reading comprehension is often an issue (Nation, Clarke, Wright & Williams, 2006). For example, a child may learn to read the written word CAT aloud, but be unable to match the written word to a picture of CAT; demonstrating a lack of comprehension as to what CAT means.

Sight word reading has been identified as an important skill for a number of reasons. Teaching sight-reading can facilitate participation in various daily living activities such as reading a grocery list and locating the corresponding items (Bowder & Lalli, 1991). In addition, because irregularities of spelling in the English language make the phonetic analysis of some words challenging, learning sight words provides an alternative to learning to decode words phonetically (Bowder & Lalli, 1991). Similarly, for individuals with disabilities, it may not be feasible to teach phonetic reading. In fact, some research has shown that some children with Autism may learn to read by memorizing features of words (i.e. shape) rather than by learning phonetics (Frith &

Snowling, 1983). As is discussed by Nation and colleagues (2006), case studies showing failure of children with Autism to acquire reading of non-words in contrast to real words (Aaron, Frantz, & Manges, 1990) support the idea that these children may be learning to read by sight rather than by decoding. What needs to be investigated, however, is how to ensure that sight word reading transfers to functional reading with comprehension, rather than remaining a rote skill with little applicability.

Stimulus equivalence has been demonstrated to be effective in teaching reading in laboratory settings, with children and adults with and without developmental disabilities (e.g., De Rose, De Souza, & Hanna 1996; Melchiori, De Souza, & De Rose, 2000; Sidman, 1971; Stromer & Mackay, 1992). “Stimulus equivalence provides methods for analyzing how physically dissimilar stimuli come to be treated as equivalent to, or substitutable for, one another in certain contexts” (Green, 2001, p.79). The recommendation for the application of stimulus equivalence technology to field settings has been repeated numerous times in the literature (e.g., Cautilli, Hancock, Thomas, & Tillman, 2002; Green, 2001; Stromer, Mackay, & Stoddard 1992). Despite these recommendations, applied studies of stimulus equivalence continue to be limited and applied stimulus equivalence studies in children with Autism are even more limited. The effectiveness of stimulus equivalence in the real world for children with Autism remains to be seen.

The purpose of this study was to investigate the field-effectiveness of stimulus equivalence technology in teaching reading skills to young children with Autism. Specifically we sought to determine if children with Autism could be taught to read and comprehend words using stimulus equivalence technology, in an applied setting.

Background

What is Autism?

Autism was first described by Kanner (1943). He described a group of children with deficits in speech and language, rigidities in routines, and difficulties with relationships. According to the Diagnostic and Statistical Manual of Mental Disorders-IV (American Psychiatric Association, 1994), Autistic Disorder is marked by signs of abnormal development apparent prior to age 3 and is characterized by delays in social skills, language, and restricted or repetitive patterns of behavior and interests. Autism is a developmental disorder, which occurs along a spectrum with a wide range of severity (Allen, 1988; Wing, 1988).

Outcomes of Interventions for Autism

Autism is considered to be a life-long disorder (Newschaffer, et al., 2007; Simpson, 2005). At this time, despite some non research-based claims to the contrary, there are no known cures for Autism (Elder, 2002; Smith, 2005). In fact, there is evidence in the literature indicating that as many as 75% of individuals with Autism, when followed through their lifespan, continue to experience poor outcomes later in life (Seltzer, Shattuck, Abbeduto & Greenberg, 2004; Smith, 2005). However, Applied Behavior Analysis (ABA), and more specifically Intensive Behavioral Intervention (IBI), has been demonstrated in the literature to lead to significantly improved outcomes for some children and can be considered a “best practice” treatment for children with Autism (Schreibman, 2000). IBI is a form of intensive early intervention, based on the principles of ABA and specifically designed for children with Autism (Maurice, Green & Luce, 2002). IBI has led to the promise of improved outcomes for children with Autism. Lovaas (1987) found 47% of children who had received IBI were indistinguishable from their peers when followed up at age 7. Similarly, McEachin, Smith, and Lovaas (1993) followed up with the same participants and found that participants remained

indistinguishable from their peers, at age 13. More recently, an investigation on the effectiveness of IBI for 332 children in Ontario was completed (Perry et al., 2008). They found that 75% of children demonstrated at least some improvement while in IBI, and 11% achieved outcomes similar to Lovaas's (1987) "best outcome" children. These results provide some evidence of the effectiveness of IBI under real world conditions. There is variability in the outcomes of children in IBI with some children making large gains and others making little to none (Perry et al., 2008). It is essential that evaluations of efficacy and effectiveness of components of IBI continue so that the components related to variability in outcome can be identified.

Stimulus Equivalence

Equivalence can be thought of as learning what words mean (Stromer et al., 1992). In its simplest form, stimulus equivalence involves the teaching of two conditional discriminations that share one stimulus (e.g., matching written word to dictated name and matching picture to dictated name) and then testing all the possible conditional discriminations not taught (Saunders & Green, 1999). Stimulus equivalence involves the formation of classes of equivalent stimuli. A stimulus class is a group of stimuli that have a common relationship (Cooper, Heron & Heward, 2007), and in the case of stimulus equivalence, the stimuli involved are arbitrary; meaning they have no physical similarities in common. The demonstration of stimulus equivalence is said to have occurred when the individual matches any member of a stimulus class with any other member of that class even if only ever trained on a subset of possible matches (Eikeseth & Smith, 1992).

Stimulus equivalence is a phenomenon that was investigated by Sidman (1971). In his study, a subject who already matched 20 pictures to dictated names and named the pictures orally, was taught to match the written words to the same dictated names. Following the training, and without any additional teaching, the learner matched the

words to corresponding pictures, the pictures to the words, and named the words orally. In this study, the dictated names, the pictures, and the printed words were all said to have formed a class of equivalent stimuli (Sidman, 1994).

An equivalence class is only said to have formed if reflexivity, symmetry, and transitivity, as described below, have been demonstrated (Sidman & Tailby, 1982). It is important to note that not all conditional relations are demonstrations of stimulus equivalence (Green, 1990).

Reflexivity. Reflexivity describes the performance of matching a stimulus to itself and can be denoted as A-A, B-B, or C-C (Sidman & Tailby, 1982). Sidman and Tailby argue for the testing of identity matching in the absence of reinforcement as proof of reflexivity. For example, being able to match (a) the written word dog to an identical version of that word, or (b) identical pictures of dog in the absence of any programmed consequences, would be considered examples of reflexivity. Sidman and Tailby also refer to reflexivity as generalized identity matching.

Symmetry. In contrast, symmetry is a demonstration of reversibility occurring without direct teaching or reinforcement (Sidman & Tailby, 1982). For example, if a subject were taught to match a picture of cat to the written word CAT, then symmetry would be demonstrated if, without any additional teaching, the subject was able to match the written word CAT to its corresponding picture. If A is related to B, then B is related to A. In sum, symmetry is demonstrated if the sample stimuli become comparison stimuli and the comparison stimuli become samples (Sidman, 2000).

Transitivity. Finally, transitivity involves the emergence of new skills that derive from earlier training (Green & Saunders, 1998; Sidman & Tailby, 1982). Driscoll and Kemp (1996), describe transitivity in terms of the subjects' proficiency with a task. More specifically, they define transitivity as the emergence of proficiency on a task for which explicit training was lacking. For example, transitivity is demonstrated if a subject is

taught to match a picture of CAT to its dictated name (A-B) and is also taught to match the written word CAT to the same dictated name (A-C), then without any additional teaching, matches the written word CAT and the picture of CAT (B-C). If A is related to B, and B is related to C, then A is related to C.

Summary. If reflexivity, symmetry, and transitivity have been demonstrated then in the examples used above, the corresponding written words, pictures, and dictated words are said to have formed classes of equivalent stimuli.

Testing for stimulus equivalence. From a practical perspective, the testing for reflexivity, symmetry, and transitivity can be conceptualized as follows. Reflexivity can be tested by conducting tests of identity matching (Sidman & Tailby, 1982). For example, when the subject can match the written word dog to the written word dog, a picture of dog to a picture of dog or imitate the spoken word “dog,” reflexivity has been demonstrated. Symmetry and transitivity can be tested using combined testing (Sidman & Tailby, 1982). According to Sidman and Tailby (1982), following the teaching of two relations with a sample stimulus in common (e.g. AB and AC) combined testing begins by testing for emergence of a new relation and the symmetry of that relation (i.e. BC and CB). As an example, after teaching a subject to match a picture of DOG to the dictated name “DOG” and to match the printed word DOG to the dictated name “DOG” testing would then follow. If the subject then, without teaching, matched the picture of DOG to the printed word DOG, and the printed word DOG to a picture of DOG, then one could conclude that equivalences classes had formed.

Challenges in testing symmetry with auditory stimuli. Sidman (1994) highlighted some challenges in the testing of symmetry when auditory stimuli are involved. Specifically, Sidman argues that when one or more of the stimuli are auditory rather than visual in nature, that a direct test of symmetry of the conditional discriminations involving the auditory stimulus is not possible. This is because auditory stimuli presented

simultaneously cannot be discriminated. For example, if A is an auditory stimulus and B and C are visual stimuli, and the individual is taught AB and AC, BA and CA cannot be tested directly. In this case, Sidman argues that symmetry can only be inferred, based on the emergence of BC and CB.

Some authors have explored this procedural challenge and have proposed alternative testing methods permitting the use of auditory stimuli (Dube, Green & Serna, 1993). Savona (unpublished, 2008) implemented a modification to allow the testing of symmetry using auditory stimuli, in typical IBI settings (i.e. without the benefit of computer-presented stimuli more common in laboratory studies). In her study, symmetry was tested using auditory stimuli, presented successively by voice recorders placed as comparisons in front of the participant. The participant could match the comparison stimuli to the sample by selecting the voice recorder, which matched the auditory stimulus. This type of modification allows the symmetrical relations to be tested (albeit indirectly) as opposed to merely inferring their emergence.

Interestingly, studies have found two additional relations to emerge that have yet to be discussed. Specifically, participants in stimulus equivalence studies have not only demonstrated written word to picture matching and picture to written word matching, but also demonstrate the emergence of picture naming and oral naming of the printed words (Sidman, 1971; Sidman & Cresson, 1973). While not a necessary criterion for equivalence (Sidman & Tailby, 1982), the possibility of the emergence of oral naming following the above teaching and testing methodology, provides even greater support for the enhanced efficiency of this approach in teaching reading skills to children with autism.

Stimulus Equivalence Research

Reading and equivalence. From a stimulus equivalence perspective, reading may be regarded as being types of stimulus-response relations where the stimuli are written

words (Sidman, 1994). According to Sidman (1994), reading includes a number of components, including: oral reading, matching printed words to pictures (and vice versa), and matching written words to dictated names. Reading can be defined as a combination of these components (Sidman, 1971). Some of these tasks, including matching written words to dictated names, are simply auditory discrimination tasks while others, such as matching written words and pictures, require comprehension (Sidman, 1994). One can argue, that all four of these components of reading are required to demonstrate functional rather than rote or memorized reading. A review of normal development reveals that children typically learn to “understand words before they learn to read them with comprehension” (Sidman, 1994, p. 24) and typically name objects and pictures before words. Further, Sidman (1994) explains that it is usually not until the first or second grades that written words come to have meaning, and for children with disabilities, reading comprehension does not always occur. Sidman also argued that the emergence of skills in reading comprehension, as opposed to only auditory comprehension, marks a critical stage in development; that must be addressed in children who do not make this transition naturally. Later, Sidman (2009) argued that someone who matches printed words to pictures understands those words and therefore demonstrate a form of reading comprehension. He therefore suggested that equivalence relations might allow the learning of reading comprehension (i.e. the emergence of skills) even in the absence of direct teaching.

Stromer and colleagues (1992) argue that the difficulties that children often encounter in learning to read sentences are related to issues occurring at the individual word level. They suggest using stimulus equivalence as a systematic approach for assessing, analyzing, and remediating the problems that children often have in naming, writing, and spelling words.

Experimental studies of stimulus equivalence. Since Sidman’s original study, many

laboratory studies of stimulus equivalence have been conducted, with adults and children with and without developmental disabilities. These studies have incorporated a stimulus equivalence paradigm to teach sight-word reading and spelling (De Rose, De Souza, & Hanna 1996; Melchiori, De Souza, & De Rose, 2000; Stromer & Mackay, 1992), math and geography skills (Hall, DeBernardis, and Reiss, 2006; LeBlanc, Miguel, Cummings, Goldsmith, & Carr, 2003; Lynch & Cuvo, 1995; Maydak, Stromer, Mackay, & Stoddard, 1995), money skills (Cuvo, Veitchm, Trace, & Knoke, 1978; Trace, Cuvo, & Criswell, 1977), name to face matching (Cowley, Green, & Braunling-McMorrow, 1992), and Greek or other arbitrary symbols (Devany, Hayes, & Nelson, 1986; Eikeseth & Smith, 1992; Green, 1990; Lazar, Davis-Lang, & Sanchez, 1984; Saunders, Wachter & Spradlin, 1988; Wetherby, Karlan & Spradlin, 1983).

Laboratory applications of stimulus equivalence methods to teach reading and spelling have demonstrated the success of this technique. In one example, Melchiori, De Souza, and De Rose (2000) taught a sample of first graders, preschoolers, special education students, and adults to match printed words to dictated words and to construct printed words. Without any additional teaching, participants matched printed words to pictures, pictures to printed words, and read the words orally. Dictated words, printed words, and pictures were said to have formed classes of equivalent stimuli. It is important to point out however that during both testing and teaching trials in this study, reinforcement, in the form of confirmation and praise, followed correct responding. It is possible that this may have impacted the results of the tests for equivalence, as the skills may have been taught directly (e.g. through reinforcement delivered contingent on correct responses) as opposed to emerging without direct teaching.

Applied studies of stimulus equivalence. Despite considerable research in the laboratory, stimulus equivalence applications in clinical settings have been limited (Vause, Martin, Marion, & Sakko, 2005). Sidman (1994) reported that he had no

knowledge of any schools that were applying the technology of stimulus equivalence in their teaching. The recommendation for the application of stimulus equivalence technology to field settings has been repeated numerous times in the literature (Cautilli, Hancock, Thomas, & Tillman, 2002; Green, 2001; Stromer, Mackay, & Stoddard, 1992). Applied studies of stimulus equivalence continue to be limited however, and applied stimulus equivalence studies in children with Autism are even more scarce (Eikeseth & Jahr, 2001). Mackay (1985) conducted a study with three teenage boys diagnosed with intellectual disabilities. The purpose of the study was to apply methods from the laboratory in desktop procedures for teaching. Mackay's subjects already matched colored patches to their dictated words and named the colors orally but did not name the printed words, match the printed words with their dictated names, or match the printed color words with the color patches. They were taught to use anagram letter cards to spell out the names of the colors. Following teaching the subjects were tested in the three previously unknown skills described above. While there was some variability amongst subjects, performance improved on the skills previously not in the participants repertoires and equivalence relations were said to have emerged. This study extended the results of Sidman's work with picture names and pictures to a different type of stimuli; color words, and color cards. The results of this study demonstrated that emergent relations were possible for some participants using tabletop rather than laboratory methods. One possible limitation of this study however, is that tokens and verbal praise were delivered to the boys throughout testing trials. While the use of reinforcement during the testing phases is consistent with other studies (LeBlanc et al., 2003; Melchiori, et al., 2000) it could result in systematic teaching rather than emergent relations.

Driscoll and Kemp (1996) conducted an applied study to teach six boys, with moderate intellectual disabilities word reading, picture labeling, and reading comprehension, using two different methods. Four of the six subjects were diagnosed

with either Down Syndrome or a developmental disability. The diagnoses for the remaining two subjects are unknown because all information regarding these subjects was dropped from the study as they failed to read the sight-words. In this study, children were taught 2 sets of words, each with a different method. The first method involved teaching oral naming of the written word and oral naming of the picture. Subjects were then tested in reading comprehension (matching word to picture). The second method involved teaching oral reading of the words and reading comprehension. Subjects were then tested in labeling of pictures. The authors conclude that some learning of untrained tasks occurred, but that acquisition of the trained skill was quite slow. This may be due to the fact that the children were unable to label the pictures prior to the intervention. Given that children typically learn to name objects and pictures before words (Sidman, 1994), teaching skills involving the written words prior to knowledge of the pictures may have created some challenges. In addition, the omission of the data for the two failed participants limits the interpretation of these findings.

In another applied study, Eikeseth and Jahr (2001) evaluated the UCLA reading and writing program. In this study, four children with Autism and three typical children were taught reading skills (defined by the authors as responding to printed instructions by matching words to pictures), writing skills (defined by the authors as labeling by matching pictures to written words), and sign language. The study examined, among other things, whether symmetry occurred more often using reading and writing than using sign language. The authors found that after teaching reading (word to picture matching), writing emerged (picture to word matching). While this was an applied study, it only tested for symmetry (i.e. trained word to picture matching, and tested picture to word matching) but provided no demonstration of reflexivity or transitivity.

Considering the limitations of the aforementioned studies and the absence of other applied research with children with Autism, one can conclude that to date there are

no published studies on the effectiveness of stimulus equivalence to teach reading skills to young children with Autism. Specifically, while the efficacy of stimulus equivalence has been well documented, its effectiveness in practical settings for children with Autism remains to be demonstrated.

Purpose

The purpose of this study was to investigate the field-effectiveness of a stimulus equivalence paradigm for teaching reading skills to young children with Autism. We want to know if children with Autism, in a typical IBI setting, could be taught to read and comprehend words using stimulus equivalence technology. This study is significant and valuable for a number of reasons. First, reading has been documented in the literature to be an essential skill for success in school and in life (Rinaldi, Sells, & McLaughlin, 1997). Second, given the long wait list and high costs for intensive behavioral intervention, any technique that has the potential to increase the efficiency of teaching should be studied. The benefit of stimulus equivalence paradigms to clinicians, however, is that they can provide practitioners with an efficient teaching methodology that gives them something for free (Green, 2001). In Sidman's original study, for example, after just 15 hours of instruction, the participant was taught to relate 20 words to their corresponding written words and without any additional teaching was able to do 40 additional tasks involving the relations between picture and text (Sidman, 1994). If stimulus equivalence can be demonstrated to be effective for teaching reading skills to children in IBI settings, then perhaps it can also be demonstrated to be useful to teach other skills. Stimulus equivalence has the potential to increase the efficiency and economy of intensive behavioral intervention by increasing the number of skills acquired in a period of time. Given the incredible number of successful demonstrations of stimulus equivalence in laboratory settings dating back to Sidman (1971), it is surprising that the

field effectiveness of this approach remains to be tested. Hence, this research is both a valuable endeavor and is long overdue.

Research Questions

This study attempted to determine if children with Autism could learn to read using a stimulus equivalence paradigm by answering the following research questions:

1. Can children with Autism demonstrate reflexivity as evidenced by accurate performance in identity-matching tasks using names, words, and pictures?
2. Can children with Autism demonstrate symmetry and transitivity (combined testing) of printed words, pictures, and dictated words, as evidenced by the emergence and reversibility of written word to picture matching, and the reversibility of matching pictures or words to their dictated names?
3. Does oral naming emerge following the teaching and emergence of the relations described above?
4. Do emergent relations and the relation taught generalize to novel stimuli not presented during the baseline or post test phases?
5. Are emergent relations maintained over short durations of time (i.e. 7-10 days)?
6. Do children with Autism acquire the taught relation (i.e. matching written words to dictated names) more quickly, following the demonstration of emergent relations on a previous word set?

Method

Participants

Participants in this study were selected from children who were already clients of the Toronto Partnership for Autism Services (TPAS). In order to be clients of TPAS,

children require a confirmed diagnosis towards the severe end of the Autism Spectrum by a qualified professional. Children were selected from three different centre-based IBI programs within the TPAS program. Children were identified by senior clinicians as appropriate participants. Four children were selected based on the following criteria: a) evidence of some sight word reading or clinical readiness to learn to read, c) ability to orally name pictures, d) ability to match pictures to dictated name and, e) clinical supervisor agreed to participation of the Research Team. Three additional participants were approached but declined participation in this study.

Participant 1. James was 7 years and 9 months old at the start of this study. He was diagnosed with Autism Spectrum Disorder at the age of 2 years and 11 months. James had been in IBI for approximately four years and was attending a centre-based IBI program for 29 hours per week, at the time of the study. James's rating on the CARS, as well as his age equivalence and general maladaptive scores from the Scales of Independent Behavior-Revised short form (SIB-R) can be found in Table 1. James' scores on both of these tests were indicative of moderate severity, with a CARS rating in the mild-moderate range and a maladaptive score on the SIB-R short form in the moderate range (CARS and SIB-R will be described in more detail in the measures section below).

Participant 2. Billy was 5 years and 1 month at the start of this study. He was diagnosed with Autism at the age of 3 years and 1 month. Billy had been in IBI for approximately eight months and was attending a centre based IBI program for 24 hours per week, at the time of the study. At the completion of the intervention phase for Word Set One, Billy transferred from his current centre to the same centre as Participant 1 (for reasons unrelated to the study). At this time, Billy's hours of service increased from 24 hours per week to 29 hours per week. Billy's rating on the CARS was in the mild-

moderate range, and his score on the SIB-R short form was indicative of maladaptive behaviour that was marginally serious (Table 1).

Participant 3. Henry was 5 years and 10 months at the start of this study. He was diagnosed with Autism at the age of 2 years and 10 months. Henry had been in IBI for approximately one and a half years and was attending a centre based IBI program for 20 hours per week. Henry also attended a school placement for 6 hours per week. Henry was transferred to a new centre based classroom after the completion of the intervention phase for Word Set One (for reasons unrelated to the study) and his IBI hours increased to 29 hours per week. Henry's rating on the CARS was in the mild-moderate range, and his score on the SIB-R short form was indicative of maladaptive behaviour that was marginally serious (Table 1).

Participant 4. Pat was 5 years and 6 months at the start of this study. He was diagnosed with Autism and a developmental disability in the mild-moderate range at the age of 2 years and 9 months. Pat had been in IBI for approximately 6 months and was attending a centre based IBI program for 25 hours per week, at the time of the study. Pat's rating on the CARS was in the mild-moderate range, and his score on the SIB-R indicated maladaptive behavior was in the normal range (Table 1).

Table 1:

Participant Description

Participant	Age	<u>CARS</u>	<u>SIB-R Short Form</u>	
		Autism Rating	General Maladaptive Index	Age Equivalent
1	7yr 9mo	Mild-Moderate	Moderate	6yr 4mo
2	5yr 1mo	Mild-Moderate	Marginal	4yr 7mo
3	5yr 10mo	Mild-Moderate	Marginal	3yr 7mo
4	5yr 6mo	Mild-Moderate	Normal	3yr 1mo

Setting and Staffing

All phases of teaching and testing occurred within the child's regular IBI centre, under the supervision of his own clinical team, with additional support and consultation from members of the research team. Sessions took place either in the child's workspace or in other areas of the centre. All teaching and testing was completed by the child's usual therapists. Members of the research team collected data on interobserver agreement and procedural integrity. Note that two of the graduate students involved in this research for course credit were the existing clinicians for two of the children who participated in this study and (with the transfer of Participants 2 and 3 to their new centre) the lead student investigator was the supervising clinician for 3 of the 4 children in the study (with the 4th participant being supervised by one of the other graduate students involved). This supported the plan for field effectiveness, as the researchers were also active clinicians with the participants. The thesis/project supervisor was also the clinical supervisor for one location, and the clinical supervisors for the other two locations agreed to participate.

The methodology was consistent with their day-to-day clinical practices as scientist practitioners and behavior analysts.

Figure 1 shows a summary of the educational background of the instructor therapists involved. Staff education ranged from having earned a college diploma, to being currently enrolled in a Master's degree. None of the staff had any additional certification or registration (i.e. BCABA, BCBA or Psychology). The instructors had a mean 1.5 years experience in their current position, and the mean number of prior years of experience was 1.3 years.

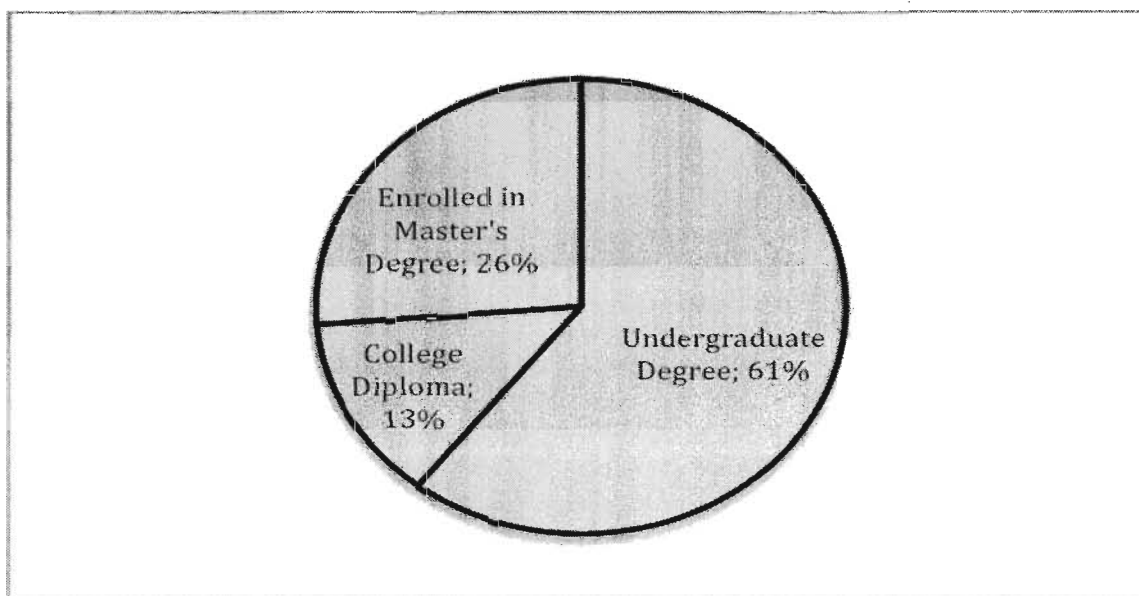


Figure 1. Educational experience of instructor therapists

Materials

The stimuli for this study were a combination of written words and pictures. Word cards were approximately 4 cm in height and 15 cm in length and included a combination of upper case and lower case letters (i.e. Juice, JUICE, juice), in 5 fonts. Picture cards were 5 cm x 5 cm in size, and included real pictures and computer drawings. Five different exemplars of each picture were provided to enhance generalization of training (Stokes & Bear, 1977) and for use in tests. A sample of the

stimuli used can be found in Appendix C. Words selected for this study were words that were functional to the child (i.e. words that may be encountered in the real world). This differs from some of the literature, where symbols or pseudo words are used as stimuli in order to control for any contamination of the study (Devany et al., 1986; Eikeseth & Smith, 1992; Green, 1990). In the present study, however, the focus was on the field effectiveness of stimulus equivalence. In Ontario, where waitlists for IBI are long and time is precious (from an early intervention perspective) teaching pseudo words would not be clinically appropriate. In order to ensure the functionality of the skills taught for each learner, the words, and pictures chosen were individualized. Four exemplars of each stimulus (with the exception of the auditory stimuli) were used during the baseline, teaching, and post test phases and one novel exemplar was used for the generalization phase.

Measures

Severity of Autism. All participants underwent an assessment of their Autism severity using the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1988). The CARS is an observation scale on which observed behaviour is recorded for 15 dimensions or symptoms. The scale yields both a total score and a categorical diagnosis. The reliability and validity of the CARS have both been documented in the literature. Specifically, the CARS has good internal consistency in both large and small sample sizes ($\alpha = .94$, $n=537$ & $\alpha = .91$, $n= 430$) (Garfin, McCallon, & Cox, 1988; Perry & Feldman, 1996; Schopler et al., 1988; Tachimori, Osada & Kurita, 2003). Inter-rater agreement and the test-retest reliability of the CARS have both been documented in the literature at high levels. Schopler and colleagues (1988) report in the manual that the inter-rater agreement ranges from .55 to .93 and the test-retest reliability is .88 ($n= 91$). Similar scores have been obtained by other researchers investigating the psychometric properties of the CARS (DiLalla & Rogers, 1994; Perry & Freeman, 1996). Furthermore,

the validity of the CARS has been demonstrated through comparisons of the CARS to the DSM-IV. Specifically, comparisons of the CARS to the DSM-IV have shown high levels of agreement in clinical diagnoses (Eaves & Milner, 1993; Perry & Freeman, 1996; Perry, Condillac, Freeman, Dunn-Geier, & Belair, 2005; Schopler et al., 1988). The CARS was administered by the lead student investigator of the research team, either alone or with another member of the research team, under the supervision of a psychologist (thesis supervisor). The investigators responsible for the CARS administration all attended full day training on the CARS led by Dr. Adrienne Perry, who has published several papers on the psychometric properties of the instrument. The lead student investigator practiced administering the CARS (with non participants) until meeting inter-rater agreement criteria with the psychologist/thesis supervisor and other research assistants who have met criteria. The other member of the research team who assisted in the CARS administrations then met inter-rater agreement criteria with the lead student investigator. Using the criteria from the Perry and colleagues (2005) study raters were required to rate 80% of items (12 of 15) within 0.5 and the total score within 4 points with an experienced rater prior to administration for the study.

Adaptive and maladaptive behaviour. The Scales of Independent Behavior-Revised (SIB-R; Bruininks, Woodcock, Weatherman, & Hill, 1996) was completed by parents of the participants. Interpretation services were offered to families whose first language was not English. The SIB-R measures functional independence and adaptive functioning levels in a variety of environments including home, educational, work, and community. The SIB-R has been normed for use with individuals from the age of 3 months to over 80 years and can be used to assess individuals with or without developmental disabilities. Studies investigating the psychometric proprieties of the SIB-R have been conducted. These studies have found reliability for the subscales to range from 0.88 to 0.98. Similarly, the item level reliability ranged from 0.7 to 0.8 and the

reliability for the maladaptive behaviour subscales are all low 0.8s (Bruininks et al, 1996).

Baseline, post tests and probe sessions. Data for all testing conditions were collected on the form provided (Appendix A). Therapists were asked to record a + if the child responded correctly and a – if the child responded incorrectly. Specific criteria for correct and incorrect responses are defined below in the task descriptions. A modified series of data sheets (Appendix A) was created midway through the study due to a revised testing order that was implemented.

Intervention data. Data were recorded for every trial using a standard data sheet already in use in one of the centres included in the study (Appendix A). Therapists recorded a √ if the child responded correctly and independently, PC if the child responded correctly when prompted and an X if the child responded incorrectly. Data were also collected on the type of prompt used (e.g., positional, gestural, physical) and the percentage of the prompt (relative to independent) on a scale of 0%-100%. Prompting data were not included in the analysis but were necessary to facilitate systematic prompt fading and errorless teaching. Mastery criterion during the intervention phase was 90% independence (for each word), across 2 days and 2 therapists.

Inter-observer Agreement. Twenty to twenty-five percent of sessions in each phase of the study was videotaped or observed live and inter-observer agreement (IOA) data was collected. IOA data was collected primarily by the members of the research team, although on a few occasions instructor therapists (rather than researchers) in the classrooms collected the data as well. The lead student researcher scored sessions with the other researchers until reliability was achieved across observers. In general, the other researchers did not score IOA independently until this reliability was achieved. On a few occasions, however, the other researcher's IOA was included from live scoring due to

some trouble viewing stimuli on the videotapes. IOA was 93% or higher across all participants and all phases of the study.

Treatment integrity. A procedural integrity checklist was created for each phase of the study and integrity data were collected through direct observation or video, for twenty to twenty-five percent of each phase, to ensure that all procedures were carried out correctly. The lead student researcher scored sessions with the other researchers until reliability was achieved across observers. The other researchers did not score independently until this reliability was achieved. Treatment integrity scores were 97% or higher for all participants, across all phases of the study.

Design

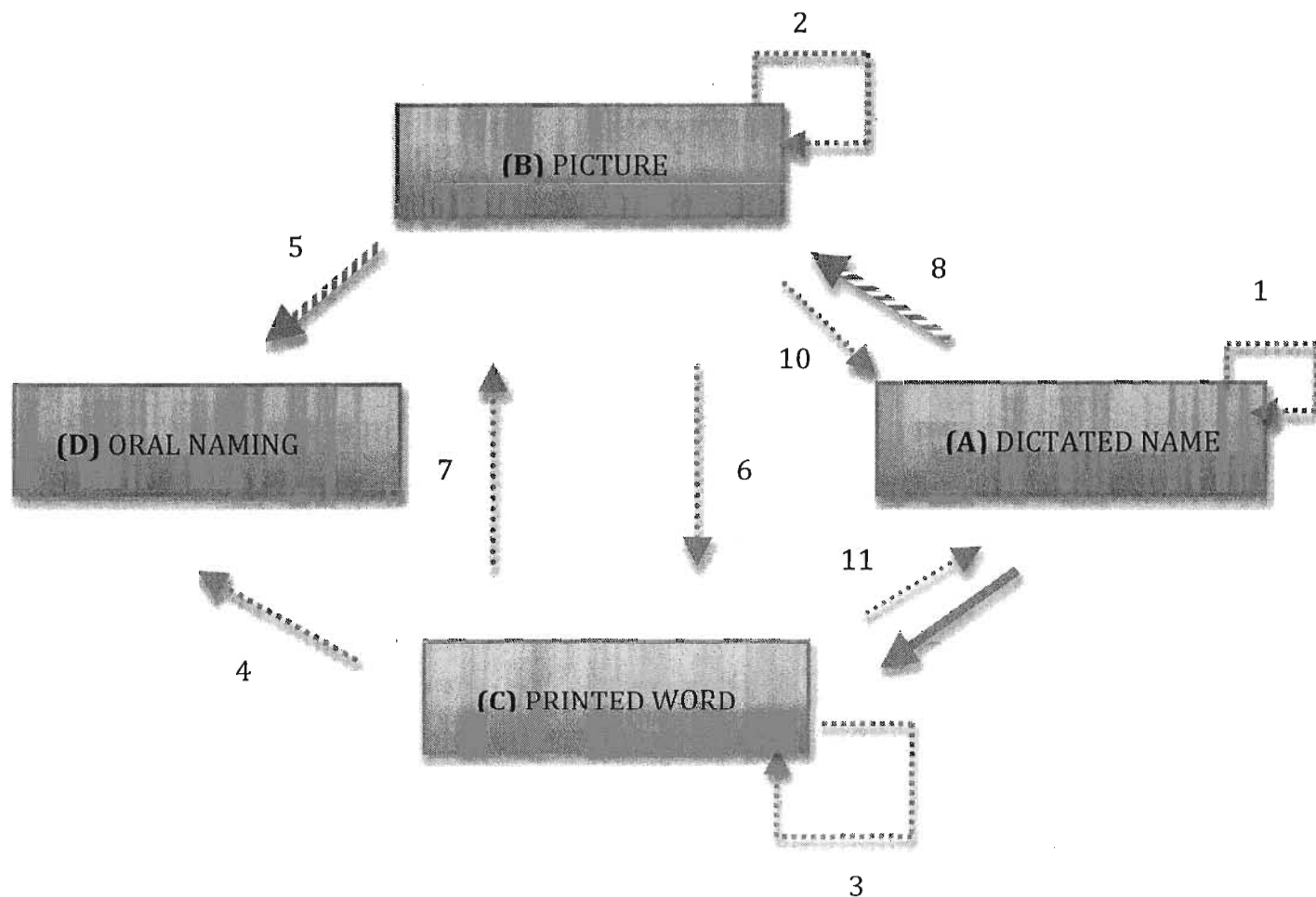
A multiple-probe design across words (Horner & Baer, 1978) was utilized for each participant in this study. According to Horner and Baer, a multiple probe design is a variation of a multiple baseline design, which allows an alternative to extended baselines. They argue that multiple probe designs can be used when extended baselines are not feasible, are reactive or when there is strong evidence of stability making extended baselines unnecessary. In a multiple probe design, intervention starts are staggered as they are within a multiple baseline. The difference, however, is that instead of collecting continuous data for the conditions remaining in baseline, the only single probes are conducted to coincide with changes in conditions (Horner & Baer). This design has been used within a stimulus equivalence paradigm (Driscoll & Kemp, 1996), providing additional support for its use in this study instead of a traditional multiple baseline design. Given the number of stimuli and the number of relations between them that required testing in this study, there was concern about the inefficiency of a multiple baseline design in the amount of time and the number of unreinforced trials that would be required. As, the purpose of this study was to demonstrate field effectiveness of stimulus

equivalence methodologies, a multiple probe design appeared more feasible and practical than a multiple baseline.

Procedure

General procedures. A combination of match-to-sample (MTS) tasks and production tasks was utilized throughout testing and training. The stimulus equivalence network that was utilized in this study is outlined in Figure 2. This network is a modification of that described by Stromer and colleagues (1992). In that paper, the authors described the oral naming tasks BD and CD (see Figure 2) as production tasks, and described tasks CB, BC, AB, and AC as MTS tasks. Each trial began with the presentation of the comparison stimuli immediately followed by the sample stimulus (e.g., for MTS trials, comparison stimuli were placed on the table and as soon as they were placed the therapist held up the sample stimulus and gave the instruction). The participant was required to either select the comparison stimuli that matched the sample stimuli, from the comparisons (by pointing to it or by placing the sample on its corresponding match) or to name the stimulus. Most of the MTS tasks (with the exception of two tasks described below) involved comparison stimuli which were presented in a messy array of six, meaning that all six stimuli were placed on the table in a randomized format. In addition, the comparison stimuli all began with the same letter and were of equal length (and where textual stimuli were involved, of the same font). The messy array was used to decrease the possibility that participants were matching based on a particular location. The similarity between the comparison stimuli was included to decrease the possibility that participants were matching based on individual features of the stimuli (i.e. the length of the word or the first letter) rather than the complete word. For trials testing the reversibility of matching pictures or words to their dictated names, an array of three comparisons was presented using a Language Master card reader to provide the auditory cues (this will be described in more detail below). This modification,

similar to that used by Savona (2008) was conducted to allow a test of symmetry for tasks of matching pictures or words to dictated names. It was not possible to conduct this test in the same manner as the others as the comparison of auditory stimuli makes the symmetry test challenging.



Task Descriptors

1. AA: Matching name to name (vocal imitation) → Test for reflexivity
2. BB: Matching picture to picture → Test for reflexivity
3. CC: Matching word to word → Test for reflexivity
4. CD: Oral naming of printed word
5. BD: Oral naming of picture → skill already in repertoire
6. BC: Matching picture to printed word → test for symmetry and transitivity
7. CB: Matching printed word to picture → test for symmetry and transitivity
8. AB: Matching picture to dictated name → skill already in repertoire
9. AC: Matching printed word to dictated name → skill to be taught
10. BA: Matching dictated name to picture → test for symmetry
11. CA: Matching dictated name to written word → test for symmetry

Figure 2. The Stimulus Equivalence Network. Striped arrows indicate a skill that is known prior to the study. Solid arrow indicates the skill taught during the intervention phase. Dotted arrows indicate skills hypothesized to emerge without formal teaching.

The following paragraphs describe the specific detailed procedures for tasks within the stimulus equivalence network.

Identity matching tasks. Matching names to names (AA), matching pictures to pictures (BB), and matching words to words (CC) can be described as identity matching. In matching names to names the therapist conducted a vocal imitation test by saying to the participant “Say_____”. A correct response was scored if the participant repeated the word presented. For both matching pictures to pictures and words to words the therapist presented a messy array of six pictures or words (all beginning with the same letter and of equal word length). The participant was handed a picture or word identical to one of the comparisons in the array and the therapist said, “Match”. A correct response was scored if the participant responded by placing the sample on the corresponding comparison. Identity matching was tested during baseline, probes, post test and follow up.

Production tasks. Oral naming of the printed word (CD) and oral naming of the picture (BD) can both be described as production tasks. In these trials, the therapist held up a single word card or picture and said to the participant “What does this say?”, “What word?” or “What’s this?”. A correct response was scored if the participant responded by reading the word or labeling the picture appropriately. Oral naming of the printed word was tested during all phases of the study. Oral naming of the picture was tested during the word selection phase and was tested as part of review during the post tests.

Matching dictated name to picture and matching dictated name to written word. Matching dictated name to picture (BA) and dictated name to written word (CA) assessed the symmetry of the AB and AC tasks of matching the pictures to dictated names and matching written words to dictated names. These skills could not be tested as directly as the other MTS tasks. To do so, a language master was utilized for testing. On these trials, the therapist ran three cards through the language master successively while holding up a picture or a written word as the sample. The therapist placed each card on the table in

front of the child and handed the child the sample stimulus and said, “match”. A correct response was scored if the participant selected the audio cue that matched the sample stimulus. Responses where the child selected the correct audio cue prior to all three cards being run were accepted as correct responses. Three trials of these tests were given in order to ensure the randomization of the auditory stimuli (i.e. target stimulus presented in the first, second, and third position). These skills were tested during baseline, the initial probe only (i.e. Word Set Two was probed when Word Set One was in baseline, Word Set Three was probed when Word Set One was in baseline but was not be probed again when Word Set Two was in baseline), and at post test.

Remaining match-to-sample tasks. Matching picture to the written word (BC), matching the written word to the picture (CB), matching the picture to the dictated name (AB), and matching the written word to the dictated name (AC) were all MTS tasks and were tested in similar ways. On these trials, a messy array of six pictures or words was presented to the participant. For the BC and CB tasks, the therapist handed the participant a printed word or picture and said, “match”. A correct response was scored if the participant matched the printed word or picture to its corresponding comparison, by placing it on the correct stimulus. For auditory-visual tasks AB and AC, the therapist delivered the instruction “Touch_____”. A correct response was scored if the participant touched the picture or word corresponding to word dictated by the therapist. All of these skills, except AB, were tested during all phases of the study. AB was tested during the word selection phase only, and was used for review during the post test phase.

Absence of reinforcement. Despite some evidence in the literature already discussed regarding the use of reinforcement during testing phases (LeBlanc, et al., 2003; Melchiori et al., 2000; Sidman, 1971), in the present study no programmed consequences followed MTS or naming trials during the testing phases. The absence of programmed consequences during testing is supported in the literature (Sidman, 1994) and is also in

line with traditional practices within IBI programming where testing in the absence of reinforcement (probe data) is conducted prior to teaching. In the present study an interspersal procedure (Horner, Day, Sprague, O'Brien & Heathfield, 1991) was utilized within the testing phases in order to maintain motivation of the participants. Interspersed requests have been highlighted in the literature as a method to increase the probability that new or difficult skills will be performed in the absence of disruptive behaviors (Horner et al., 1991). Skills that had been mastered outside of the context of this study (e.g. "clap hands") were interspersed with testing trials and these skills, or the child's attending skills (i.e. sitting nicely, etc) were reinforced during the testing phases. The frequency of interspersed tasks and the subsequent schedule of reinforcement were individualized for each participant (and was based on their current schedule of reinforcement in IBI). During training phases, the correct responses were followed by programmed consequences, consistent with the existing reinforcement procedures in place for each child.

Staff training. Prior to the start of word selection and baseline, all staff involved in the study participated in a 1.5-hour training session conducted by the three graduate students conducting the research. Training involved an overview of stimulus equivalence, an in-depth review of the teaching procedures (including prompting and error correction procedures), and testing procedures, and a review of the data collection methods. Training included lectures as well as modeling and role-playing of procedures.

Procedure Steps. The following is a list of the order of events throughout the study:

1. Supervising therapist (ST) and Clinical Supervisor (CS) identified appropriate clients for the study

2. Consent to participate in the study was sought from the instructor therapists (IT) on the IBI team. If the ITs chose not to participate the family was not contacted.
3. If the ITs consented to participate then the supervising therapist sent the family a letter of information and a consent form. A member of the SPC research team followed up with a phone call to answer any questions and determine if the family wished to participate.
4. If the family chose not to consent the ST informed them that the intervention may still be implemented without formal participation in the study if clinically appropriate.
5. If the family consented to participate in the study, the SIB-R was sent home for parental completion. In one case an interpreter was needed so the SIB-R was completed with the family, the lead student researcher and an interpreter at the centre.
6. The CARS was then completed by the lead student researcher, under the supervision of the psychologist.
7. Word selection then began with the ST, CS and IT using the procedures outlined above.
8. Baseline for Word Set One and single probes for the other sets then occurred, following procedures outlined above.
9. Teaching phase of Word Set One began, following a stable baseline, using the procedures outlined above.
10. Equivalence tests of Word Set One occurred, when the teaching step reaches mastery, using the procedures outlined above.
11. Generalization and follow-up probes then occurred using the procedures outlined above.

12. Steps 8-11 were repeated for Word Sets Two and Three.

Word selection. Each child's senior therapist or an instructor therapist familiar with the child selected 20 words that were relevant to the individual child from a list provided by the research team. All words selected were actual words (as opposed to pseudo words) that have a likelihood of being relevant to the child and that the child could name verbally and identify (i.e. match picture to dictated name) the corresponding picture.

The 20 words chosen were then assessed until three sets of three words were identified. Each word was probed three times, for each task and the accuracy, of each task, was reported as a percentage of three (0%, 33%, 66%, 100%). The order of tasks listed in the Stimulus Equivalence Network (Figure 2) was the order in which tasks were assessed in order to determine their inclusion in the study (i.e. oral naming of printed word (CD), oral naming of picture (BD), matching picture to printed word (BC), matching printed word to picture (CB), matching picture to dictated name (AB), and lastly, matching printed word to dictated name (AC)). Matching name to name (AA), matching picture to picture (BB), matching word to word (CC), matching dictated name to picture (BA), and matching dictated name to written word (CA) were not included in the word selection phase. If the child failed to name the picture or match the picture to its dictated name (BD and AB) then that word was discarded from the list. Within clinical practice it is important to be able to implement programs quickly. Extending baselines to wait for stability is not always feasible. Therefore, in this study any words for which the initial probes for the remaining relations (AC, CD, CB, BC) were not 0% were also discarded as target words. A list of the excluded words was given to the clinical team for implementation at the conclusion of the study. Probe sessions continued until 9 words were identified for which performance for the AC, CD, CB and BC tasks was 0% and performance the AB and BD tasks was 100%. The accuracy of the identity matching

relations (AA, BB, and CC) were not required to be at 0% in order for the word to be included (which is why they were not tested until baseline). Given the likely learning history of the participants with respect to identity matching tasks, it is possible that these skills were established prior to teaching. While this does create a challenge in describing reflexivity as an emergent skill, there is evidence in the literature of equivalence studies in which identity matching was already established prior to teaching (Sidman, Cresson, & Willson-Morris, 1974).

Once the sets of target words were identified for each child, the child's current treatment team was asked to avoid using these words and pictures in any other activities or teaching programs for the duration of the study. Naturally occurring contact with the target words during other times, however, was not controlled.

Table 2 shows the nine words selected for each participant. For Participant 1, word selection required testing a total of 24 words, across 5 days and a total of 447 trials. Participant 2's word selection phase required a total of 16 words to be tested, across 2 days and a total of 232 trials. A total of 18 words were tested, across 3 days and a total of 255 trials, in order to select nine words for Participant 3. Lastly, participant 4's word selection phase required a total of 27 words to be tested, across 3 days and a total of 323 trials. Appendix B, Tables B2, B4, B6, and B8, show the data for the participants' word selections. During word selection (the first phase of the study), the therapists made some initial procedural errors, in that part of the test sequence was skipped. Any words for which errors in the methodology were made were not included in the study.

Table 2:

Words Selected For Each Participant

	Word Set 1	Word Set 2	Word Set 3
Participant 1	Fork	Two	Cat
	Baby	Bike	Wolf
	Bed	Five	Book
Participant 2	Ant	Fish	Milk
	Cup	Owl	Star
	Dog	Door	Lion
Participant 3	Bike	Bus	Duck
	Glue	Fork	Cat
	Pig	Bear	Book
Participant 4	Lion	Tree	Bike
	Two	Cat	Five
	Book	Boat	Baby

Baseline

Word sets were introduced in the following manner. Word Set One was used in tests for relations AA, BB, CC, CD, BC, CB, AC, BA, and CA until a stable baseline was established, across a minimum of three data points, for each of the tasks. AB and BD, which involved skills established as known during word selection were not included. Concurrently, Word Sets Two and Three underwent a single probe session for each of the relations. In order to limit the number of tests for each participant, BA and CA relations were only tested during the initial probe, baseline and post tests (i.e. when Word Set Two entered baseline BA and CA were omitted from Word Set Three's single probe). The intervention phase (described below) began for Word Set One, after a stable baseline had been established (i.e. minimum of three data points on stable or downward trend). When Word Set One had met the mastery criteria, post tests (described below) were conducted

and then the above sequence was repeated for Word Set Two (while Word Set Three received single probes for all relations except BA and CA). This sequence was then repeated for the final word set. No programmed consequences were delivered for correct or incorrect responding during baseline. Previously learned skills interspersed with target trials and readiness skills (i.e. appropriate sitting, looking, etc) were reinforced during baseline.

Initial test order. Baseline probes for Word Set One (and the corresponding single probes for Word Sets Two and Three) for Participants 1-4 were conducted in the following order. Participants were first tested in identity matching tasks (AA, BB and CC), followed by oral naming of written words (CD). This was followed by the combined tests for symmetry and transitivity, matching pictures to words, and words to pictures (BC and CB). Next, participants were tested for the relation that was going to be taught, matching written words to dictated names (AC) and finally, symmetry of both AB and AC was tested by testing matching dictated names to pictures (BA) and matching dictated names written words (CA). Baseline for Participant 1's Word Set Two was also assessed in this order.

Revised test order. Following the post test for the first word set for Participant 1 (Word Set One) revisions to the baseline test order were made. The test order was revised by placing the identity matching tasks (AA, BB, and CC) at the end of the testing session (following BA and CA). Green and Saunders (1998) suggest that the reflexivity tests might disrupt the results of equivalence tests and as such should not precede other tests. Following the initial failures of participant 1 and 4 during the post test (which will be discussed later) reflexivity testing followed the testing of other relations to decrease the possibility of interference. Participants 2, 3, and 4 followed this revised test order for everything except baseline for Word Set One and the initial probes for Word Sets Two

and Three. Participant 1 followed this test order for all testing after the baseline for Word Set Two and second single probe for Word Set Three.

Baseline data points for tasks AA, BB, CC, BC, CB, BD, CD, AC, and AB, represented the accuracy of performance across the three words in the set. Two trials were run for each word per data point (i.e. data point one, for example represented two trials for cat, dog and bird. If the child got two responses correct out of six, 33% was recorded for that data point). Data points for CA and BA represented the combined data across the three words in the set as well, however for these tasks three trials were run for each word, per data point. Each probe session data point also represented two trials, per word collapsed across words.

Intervention

Set up. Intervention involved the child to match the written words to their dictated names (AC). The comparison stimuli were presented in a messy array of 6 where the distracter stimuli all began with the same letter and were of similar length.

Instructions. Each teaching session consisted of 10 trials (per word), completed within a single block of time or given intermittently throughout the day.

Stimulus preference testing and reinforcement procedures. Stimulus preference assessments were completed prior to intervention in order to select reinforcers. A multiple stimulus without replacement preference assessment (DeLeon & Iwata, 1996) was completed for each participant. A number of items were presented and the order in which the participant selected the items was recorded. Once an item was selected it was not returned to the array. Hierarchies of seven to nine items were identified as reinforcers for each participant. These items were available during all subsequent teaching and testing phases. Motivational systems (token systems, etc) and schedules of reinforcement currently in place for each participant's programs were utilized throughout the study. All four participants had token systems in place at the start of the study. These systems were

used to reinforce readiness and attending skills as well as skills interspersed with the target responses.

Prompting. Errorless teaching strategies were employed throughout the intervention phase by using antecedent prompts (e.g., positional, gestural, partial physical). Prompts were faded using a systematic fading procedure appropriate to each learner's history. Details on the specific prompts used for each participant are described in the results section.

Error correction procedure. If the child responded incorrectly on any given trial, an error correction procedure was followed. In the event of an error, the instructor paused and cleared the stimuli from the table. Stimuli were then represented and a prompt was delivered immediately. Following the child's response neutral feedback was delivered (i.e. "that's dog") and the instructor again paused for a minimum of 3 seconds before moving onto the next trial. The next time the stimulus (on which the error occurred) was presented an antecedent prompt, one level higher than the unsuccessful prompt, was used (i.e. if the error occurred using a 75% physical prompt, then the next trial would use a 100% physical prompt). Data were collected only on the trial of the error and not for the correction trial. Data collection began again on the next presentation of the stimulus. The specific prompt was dependent on the individual learning history for each child.

Mastery criteria. Mastery criteria for the intervention phase was 90% correct, across 2 days and two instructors (for each word), where the first trial of the day was correct.

Post Tests

Once the first word set had reached mastery criterion, the tests administered during baseline were re-administered in order to assess for the properties of equivalence (reflexivity, symmetry, transitivity) as well as the emergence of oral naming. This was broken down into five parts, where each part began with a review of the associated skills

followed by the testing for the emergence of one or two skills. A minimum of three sessions were run for each part before moving onto the next Part. Part 1, the combined test for symmetry and transitivity, began with a review of matching pictures to dictated names (AB) and matching written words to dictated names (AC). The combined test then followed with testing in matching words to pictures (CB) and pictures to words (BC). Part 2, the test for the emergence of oral naming of written words (CD), began with a review of oral naming of pictures (BD) and matching pictures to words (BC). This was then followed by testing oral naming of written words (CD). Part 3, which tested for symmetry of the written words and dictated names, began with a review of matching written words to dictated names (AC). Testing matching of dictated names to written words (CA) then followed. Similarly, Part 4, testing the symmetry of the pictures and dictated names, began with a test for matching pictures to dictated names (AB), which was then followed with matching dictated names to pictures (BA). Lastly, Part 5, the reflexivity tests, were tested in the following order, matching names to names (AA), then matching pictures to pictures (BB), and finally matching words to words (CC). Following the equivalence tests for Word Set One, the sequence already described in baseline commenced again (i.e. Word Set Two entered baseline, Word Set Three had a single probe trial for each step). This was repeated until all three words completed baseline, intervention, and post-tests. Three data points were gathered in the same manner in which baseline tests were administered (i.e. Each data point represented the combined data across the three words in the set, with 2 trials being run for each word per data point).

The sequence described above actually represents a revision to what was originally implemented. Post tests for Participant 1 (Word Set One) and for the initial sessions of Word Set One for Participant 4 were presented in a different order. Participants were first tested in identity matching tasks (AA, BB, and CC), followed by oral naming of written words (CD). This was followed by matching pictures to words,

words to pictures (BC and CB), and matching written words to dictated names (AC). Lastly, participants were tested for the matching of dictated names to pictures (BA) and the matching of dictated names to written words (CA). The order was revised when Participants 1 and 4 demonstrated inconsistent results. Participants 2 and 3 both started their post tests with the revised test order and never encountered the original testing sequence.

Generalization

Following the post test, a generalization probe was conducted, in which a single probe (as already described) was administered for CD, BC, CB, and AC. During the generalization probe, novel materials that had not been used earlier in the study were presented in order to determine if the responses are able to generalize to new stimuli. For example, if the sample stimulus was CAT, then throughout teaching and testing phases, four different pictures of CAT and CAT written in four different fonts would have been used. During the generalization probe, a fifth picture of CAT and a fifth font were introduced.

Follow Up

Follow up probes were conducted to measure the maintenance of the trained task and the emergent performances. Follow up probes for the word set that already completed the sequence (i.e. Word Set One) was completed when the subsequent word set (i.e. Word Set two) was in acquisition. Since Word Set Three was the final set to be conducted, a follow-up probe was conducted for this word set ten days after the generalization probes (described below). Thus, each word set received one follow up probe session for CD, BC, CB, AC, BA, CA, AA, BB and CC.

Error analysis. In the event that the child's performance during any phase of the study required further analysis (e.g., increasing data trend during baseline, incorrect performances on post test) this analysis was conducted by a member of the research team.

No formal data sheets were created for this purpose but rather the researcher observing the session would record the data as necessary (some examples include data regarding position preferences, a preference for a particular stimulus, errors with a specific exemplar, etc).

Procedural Modifications for Individual Participants

Participant 1. During the word selection phase, challenges occurred in identifying words for which James's performance met the criteria of 0 % for all the relations except oral naming of picture (BD) and matching the picture to its dictated name (AB) which were 100%. Observation and informal analysis indicated that James was selecting the correct response on some trials due to what appeared to be chance responding. Initially, three trials were run for each relation, for each word. In James's case, in the instance where a correct response occurred for a relation where criterion was 0%, or an incorrect response occurred for a relation where criterion was 100, the entire test sequence was repeated for the word. Words were then included if there was consistent responding within that relation (e.g., all correct or all incorrect as required).

During the baseline phase for Word Set One, James's performance appeared to indicate knowledge of the words, which were not in his repertoire during word selection. A modification was made to the procedure such that non-target words (i.e. comparison stimuli became sample stimuli) were also presented to James as target words (in the absence of data collection). For example, if the target word was dog, then James was asked to (for example) match the word dog to an identical word dog, match the word den to an identical word den and match the word doe to the identical word doe. Target and non-target words were tested for each of the relations where the initial data collected was not on a stable trend, with a minimum of two non-target word trials per target trial.

As was already discussed, James's post test order was revised following during the post tests for the first word set. The five part post test, as was already discussed was implemented and Word Set One was retested using that order.

During the baseline phase for Word Set Two, data for the CA tests, matching dictated names to written words was on a clear upward trend suggesting that learning was taking place. Following observation to confirm that the therapists were not inadvertently reinforcing his responses, an error analysis was conducted by collecting data on the participant's responses. These data indicated a clear and consistent pattern of responding. Across 13 trials, James selected comparison stimuli by first selecting the stimulus on the left, then middle, then right. This pattern was maintained across 13 trials. Instructors therefore began to place the correct comparison stimuli in a pattern that differed from James's sequence (i.e. if James had selected the middle position last, then the correct comparison would be placed in either the left or middle position).

Participant 2. During the baseline condition for Word Set One, Billy's performance on matching dictated names to written words appeared as though learning was taking place. After careful observation and analysis to confirm that reinforcement was not being provided during these trials, closer examination indicated that faulty matching was occurring. Billy was matching the words to the audio sounds that corresponded to the picture cards tested on the previous task. For example, the picture ANT was matched to the audio cue "ANT" in the test for matching dictated names to pictures. Then when presented with words to match, Billy matched the word to the audio cue "ANT" regardless of the word presented. To counteract this effect, Billy was tested on two trials of matching dictated name to picture for a non target word immediately preceding trials for matching dictate name to written word for the target word (i.e. matching picture of ARK to audio ARK, matching picture of ART to audio ART, then matching word ANT to audio "ANT").

During post-test part three (AC-CA) for Word Set One, Billy was observed to be placing his response in the middle position on every trial. A non-contingent reinforcement (NCR) system was implemented where Billy was told how many questions were to be answered and following each response a block was dropped in a cup (to signal reinforcement) and Billy was told “that’s one,” “that’s two,” regardless of whether his performance was correct or incorrect. When this procedure was ineffective, reinforcement of the trained trials (AC) was implemented followed by the test trials (CA) with the NCR still in place. This procedure was also ineffective. Finally, the array was changed from the three language master cards being placed in front of him to the cards being placed in three corners of the desk (with the reinforcement of trained trials and the NCR still in place). This change appeared to increase Billy’s attending initially, however performance quickly returned to the same pattern.

Participant 3. During post test part 3 (AC-CA) for Word Set One, Henry also struggled with the language master task. A number of revisions were implemented. To begin with, Henry was given a choice of which word he wanted to do (i.e. G word, P word, or B word) and was allowed to hold the target stimulus while the comparisons were being run through the language master. Next, a non-contingent reinforcement system similar to that already described for Participant 2 was implemented where feedback (“that’s one,” “that’s two”) paired with delivery a block dropped in a cup, was delivered following each response. Lastly, the session was repeated this time reinforcing the trained relations (AC) followed by unreinforced test trials with the NCR still in place. Similar challenges were noted for Word Set Two. During Word Set Three, an additional modification was made where Henry was allowed to run the language master cards through himself and place them anywhere on the table. Following post test for Word Set Three, post tests for both of these relations were repeated for Word Set One and Two allowing Henry to run the cards himself.

Participant 4. During the intervention phase for Word Set One, gestural prompts were demonstrated to be an ineffective method for Pat. Informal analysis indicated that as prompts were faded incorrect responding increased. Errors also were observed with the use of positional prompts, and physical prompts were not used for Pat due to a history of prompt dependency with this type of prompting. A simultaneous prompting procedure was implemented with this participant in order to increase his attending to the materials and to allow prompts to be faded. Within this procedure three types of prompts were used simultaneously, gestural prompts, positional prompts, and a verbal prompt referred to a spelling prompt. Trials began with the instruction, “touch lion” (for example) and were immediately followed by a spelling prompt L, I, O, N and a gestural prompt (instructor pointed to each letter as the word was spelled). Initially a positional prompt was used as well; the correct word was immediately in front of the participant. For the purposes of data collection, a 100% prompt was defined as the use of all three of the prompts described above, a 75% prompt was defined as the use of two of the prompts, 50% prompt was defined as the verbal or spelling prompt only, and 25% and 5% prompts were partial verbal/spelling prompts (i.e. L, I, O or L, I).

During the initial post test for Participant 4, observation revealed that Pat was choosing an incorrect comparison stimulus while looking at (and often spelling out loud) the correct comparison or was engaging in other inappropriate behaviors such as flipping the stimuli or choosing two stimuli. It was hypothesized that Pat’s behavior was due to the absence of any programmed consequences following the test trials and that the interspersal procedure was not sufficient to maintain his behavior given his current reinforcement history of receiving feedback following responses. Consequently, a procedure was introduced in which feedback was given following every response, regardless of correct or incorrect responding. This procedure involved the use of a peg board with six pegs. Prior to starting work Pat was told he had to answer six questions

and then he could have access to a preferred item or a short gross-motor break. Following each question the therapist said, “that’s one or that’s question number one” and deliver a peg to Pat to insert into the peg board. When the problem behaviors continued to occur despite this revision, and when the inappropriate behaviors were starting to transfer to other programs Pat was working on, the decision was made to correct the inappropriate behaviors as well. If Pat responded by touching two cards or flipping them inappropriately he was given feedback on his responding (but not the response) and the trial was represented. This type of feedback was used for any type of inappropriate behavior that occurred throughout testing. This feedback was given regardless of whether he touched the correct stimulus or not in order to limit the possibility that his response was being reinforced or corrected.

Results

Figures 3-6 display the summary results for each participant. The darker bars of the graph represent the baseline scores for each relation, across all three words sets. The lighter bars represent the post test scores for each relation, collapsed across all three word sets (Note: in cases where a post test was skipped due to mastery of AC during baseline the lighter bars represent the data by the end of baseline, whereas the darker bars represent the baseline at the initial probe). In situations where additional trials were required to attain stable data, the last three data points were used in constructing Figures 3-6.

Participants 1, 2, and 3 (Figures 3, 4, and 5) passed the majority of tests for equivalence for all three word sets. All three of these participants passed the tests for reflexivity as demonstrated by their completing identity-matching tasks using names, words and pictures (AA, BB, CC). They also all passed the tests of equivalence as evidenced by the emergence (combined test of symmetry and transitivity) and reversibility of written word to picture matching (BC and CB), and Participants 1 and 3

also passed the additional test of symmetry, as demonstrated by the reversibility of matching pictures or words to their dictated names (BA and CA). Participants 2 did not demonstrate the BA and CA relations. In addition, for Participants 1, 2, and 3, oral naming of written words (reading) emerged for the majority of the words.

Participant 4 passed the tests for reflexivity, as demonstrated by completion of identity-matching tasks using names, words and pictures (AA, BB, CC). He also demonstrated equivalence (combined test for symmetry and transitivity) as evidenced by the emergence and reversibility of written word to picture matching (BC and CB). He did not, however, demonstrate the reversibility of matching pictures or written words to their dictated names (BA and CA) and did not demonstrate oral naming of the written words. Participant 4's participation in the study was discontinued after the first word set. Therefore, while the summary graphs for participants 1-3 represent data across all three word sets, participant 4's data (Figure 6) is only representative of one word set.

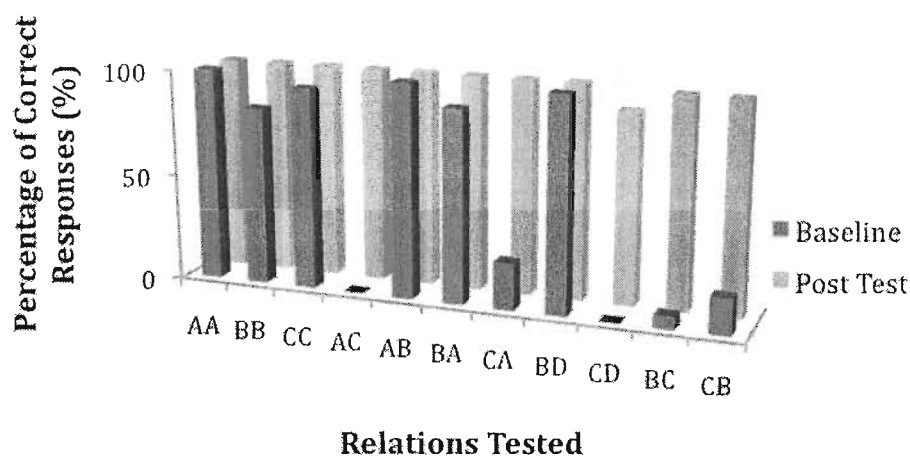


Figure 3. Summary graph for participant 1

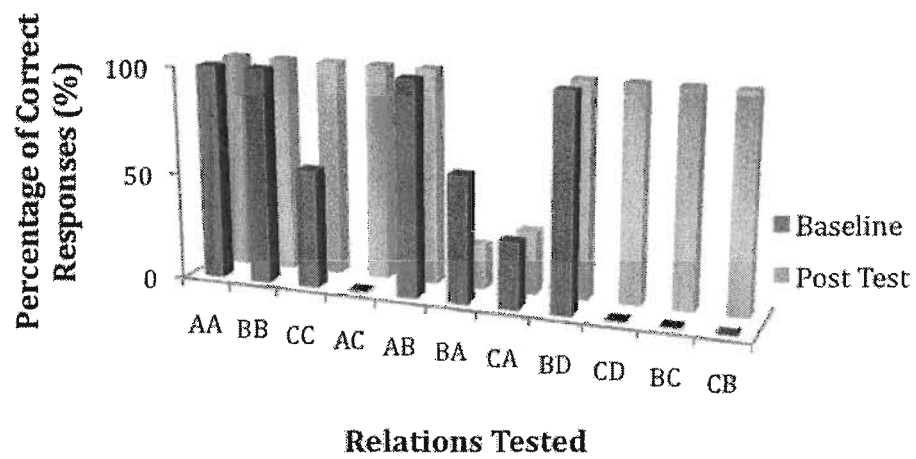


Figure 4. Summary graph for participant 2

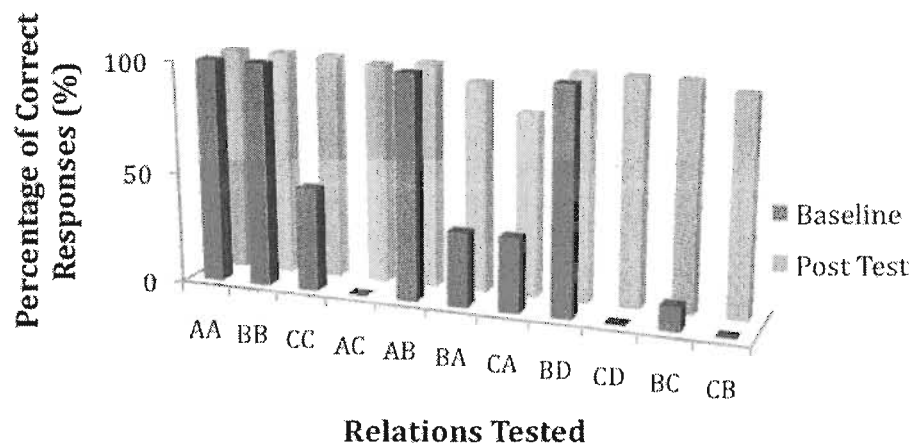


Figure 5. Summary graph for participant 3

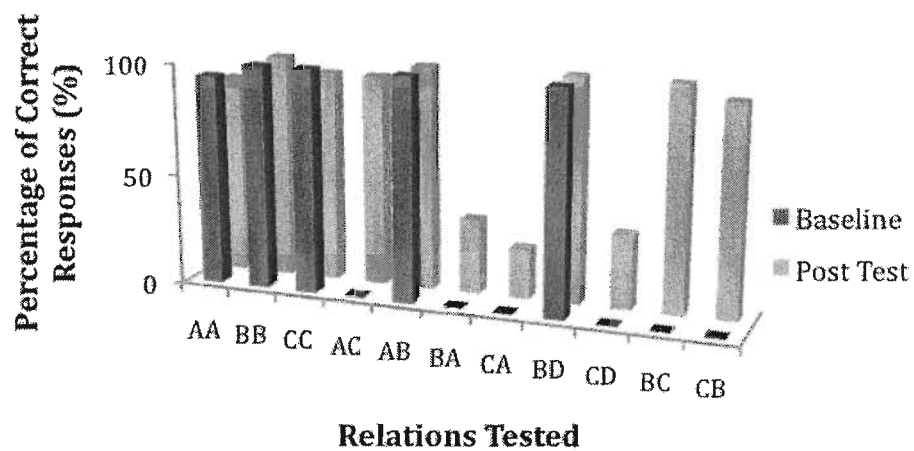


Figure 6. Summary graph for participant 4

Figure 7, which represents a summary of the generalization data for all four participants, shows that the majority of relations were generalized. All four participants generalized the matching of written words to dictated names (AC), the matching of pictures to written words (BC), and the matching of written words to pictures (BC), to novel stimuli for the majority of words (participant 2 did show some challenges in generalization for the AC and CB for Word Set Two). Participant 1, 2, and 3 generalized oral naming of written words (CD) to novel stimuli.

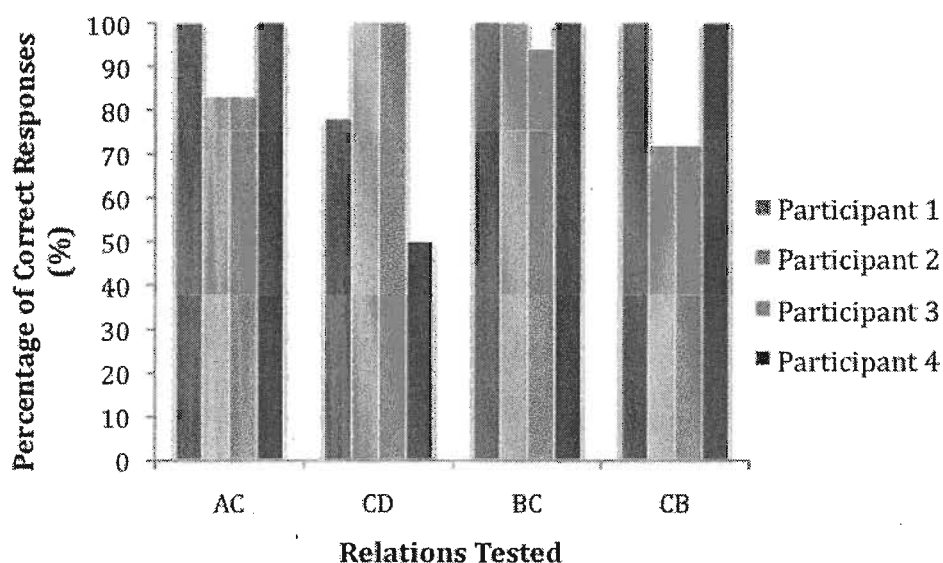


Figure 7. Generalization summary graph for participants 1-4

Figure 8 represents a summary of the follow up data for all four participants and shows that in general the relations were maintained at the follow up phase. One exception to this is the oral naming of written word (CD) for Participant 1; this will be discussed in more detail during the individual participant descriptions to follow. Lower percentages of correct responses can also be seen for the matching of dictated name to picture (BA) and the matching of dictated name to written word (CA), however these low scores are consistent with the participants' performance during the post test phase. Therefore, generally speaking, if the participant demonstrated emergent relations, then these relations were maintained at the follow up phase.

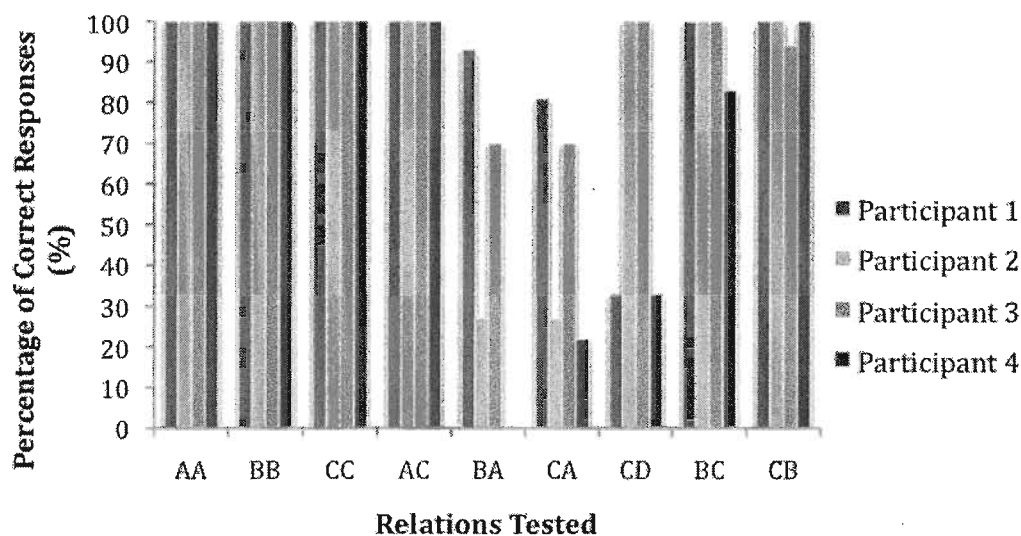


Figure 8. Follow-up summary graph for participants 1-4

Detailed results for each participant are outlined below.

Participant 1

Reflexivity. Table 3 shows the initial baseline, post test and follow up results for James's identity matching tasks (AA, BB and CC). James's performance on the initial baseline probes, across all three word sets, for matching names to names (AA) was 100%. For matching pictures to pictures (BB) performance ranged from 66%-83% for the initial baseline probes. For matching words to words (CC) James's performance on the initial baseline probes was 83% for the first and second word sets, and 100% for the third word set. For post test and follow up phases, James's accuracy on all reflexivity tests was 100%.

Table 3:

Participant 1: Reflexivity scores for all three word sets

Word Set	<u>Names to Names (AA)</u>			<u>Pictures to Pictures (BB)</u>			<u>Words to Words (CC)</u>		
	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up
1	100%	100%	100%	83%	100%	100%	83%	100%	100%
2	100%	100%	100%	66%	100%	100%	83%	100%	100%
3	100%	100%	100%	83%	100%	100%	100%	100%	100%

Word Set One. From the nine words selected for James, baby, fork, and bed were denoted as Word Set One. As was previously described, Word Set One entered baseline and at the same time a single probe session was completed for each of Word Sets Two and Three. Results for James's Word Set One for each phase of the study will be discussed next, and the results of the single probe sessions for Word Sets Two and Three will be discussed later when presenting the results for these word sets.

In order to obtain a stable baseline for James across all the baseline tasks, 417 test trials were completed. James's results of matching written word to dictated name (AC) are shown in the upper graph of Figure 9. James's accuracy was initially variable but stabilized at 0%. Figures 10 and 11 show that James's accuracy on matching dictated name to picture (BA) was initially low but quickly increased and stabilized at 89%. In contrast, his accuracy on matching dictated name to written word (CA) remained low with performances at 33%. The results of oral naming of the written word (CD), matching the picture to the written word (BC), and matching the written word to the picture (CB) are shown in the upper graphs in Figures 12, 13, and 14 respectively. James's accuracy in oral naming (CD) and in matching the written word to the picture (CB) was 0%. James's accuracy in matching the picture to the word (BC) was initially variable but eventually also stabilized at 0%.

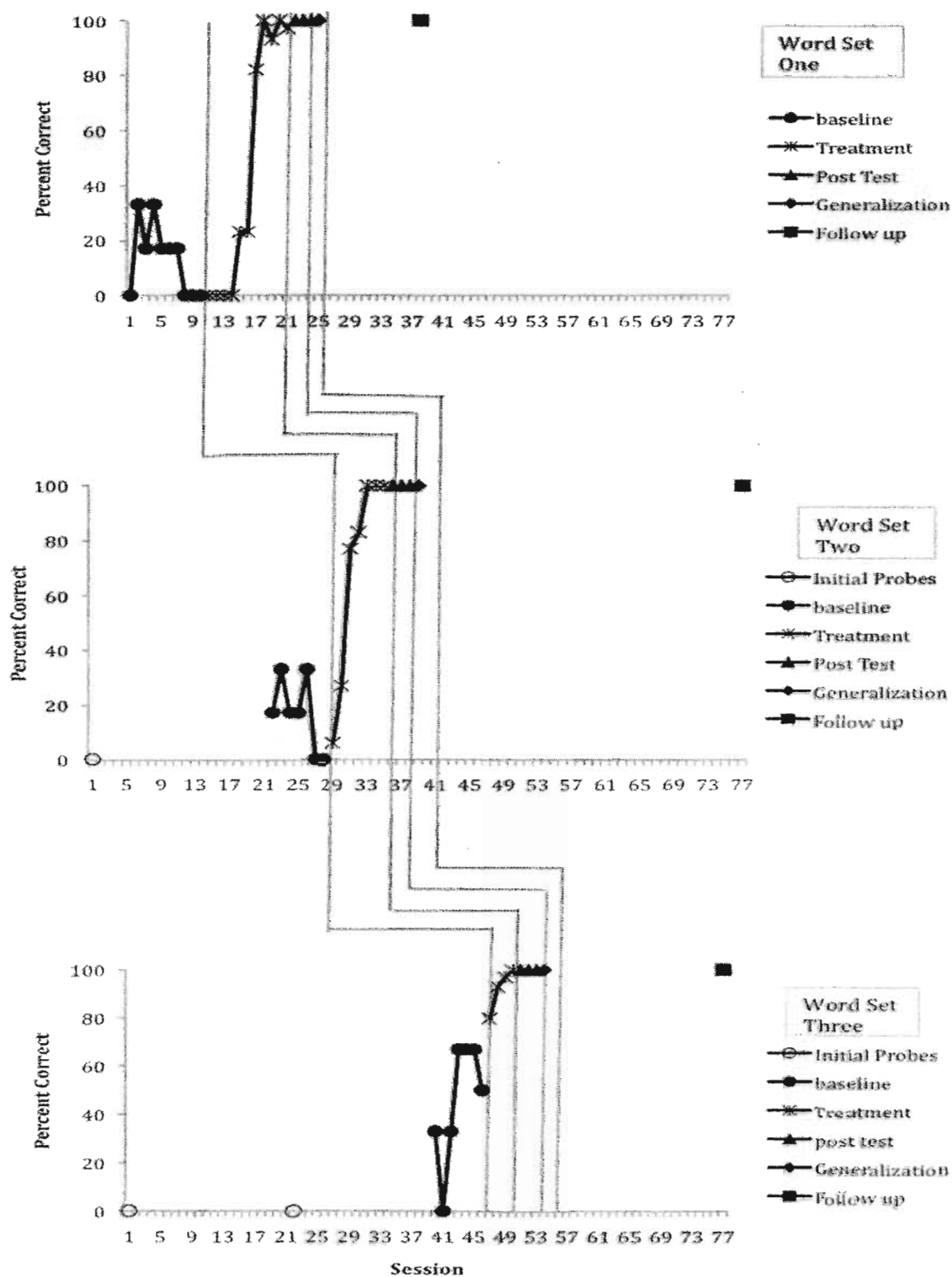


Figure 9. Participant 1: Matching written word to dictated name across all three word sets

(AC-Taught relation)

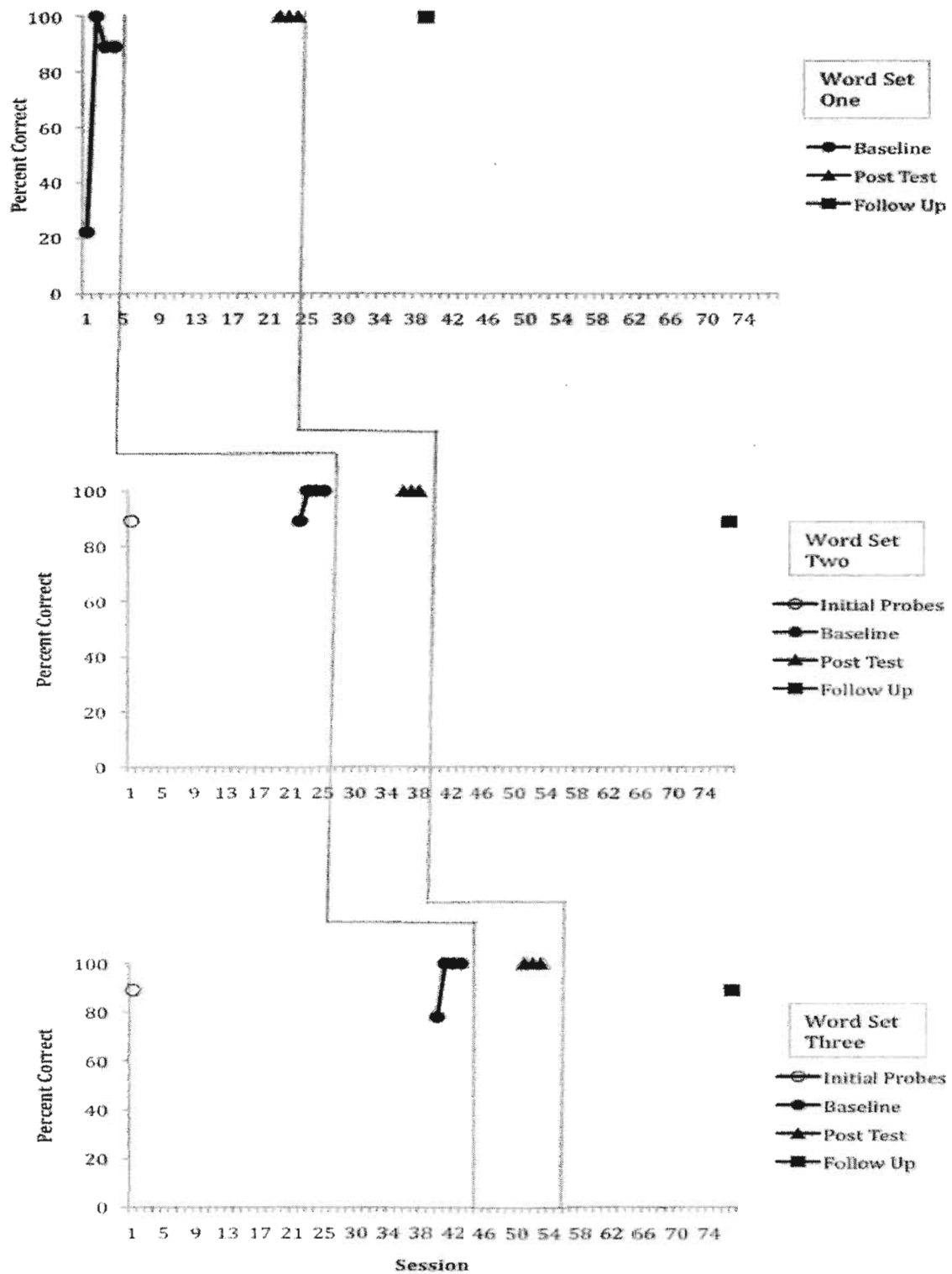


Figure 10. Participant 1: Matching dictated name to picture across all three word sets

(BA)

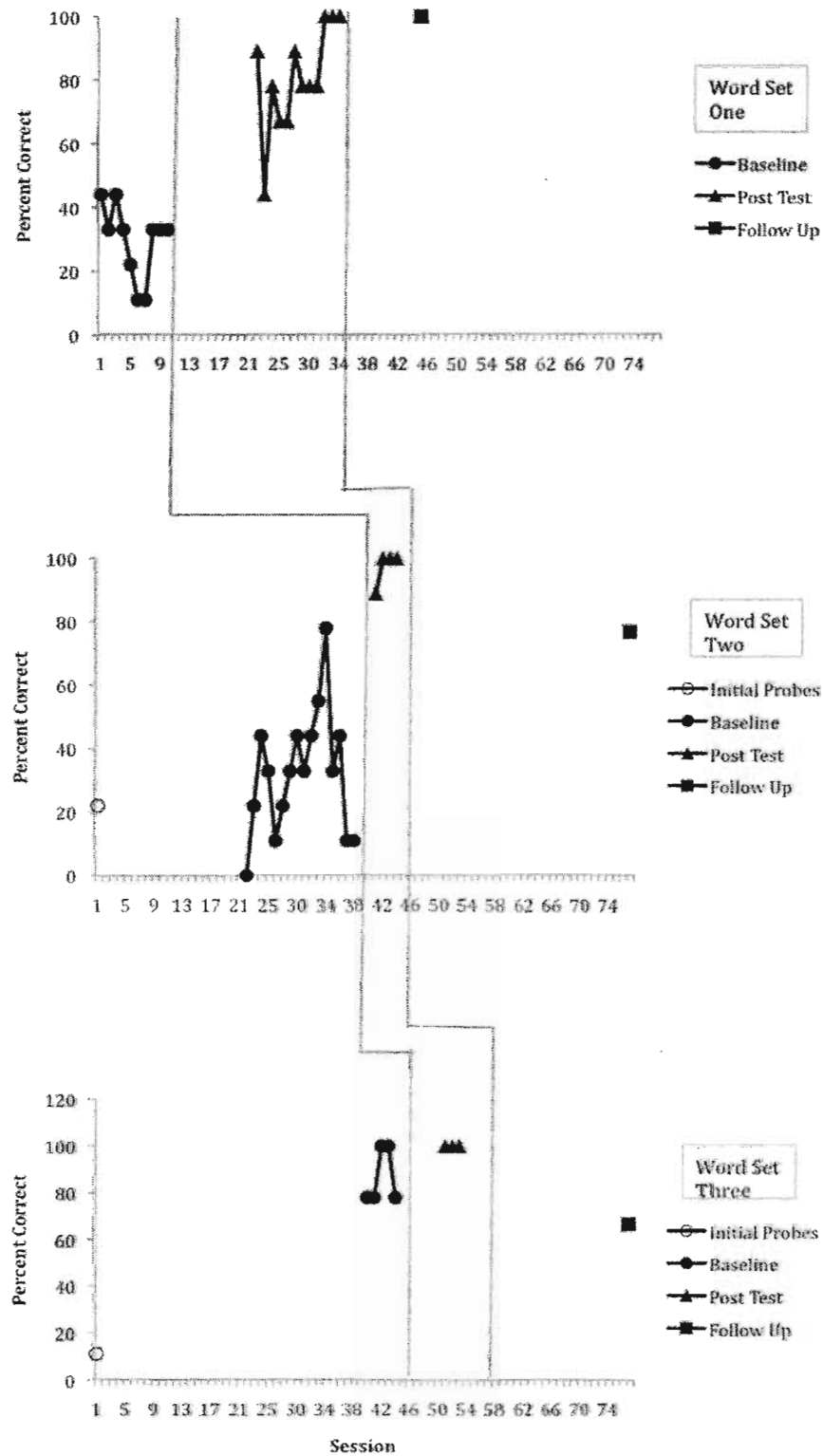


Figure 11. Participant 1: Matching dictated name to written word across all three word sets (CA)

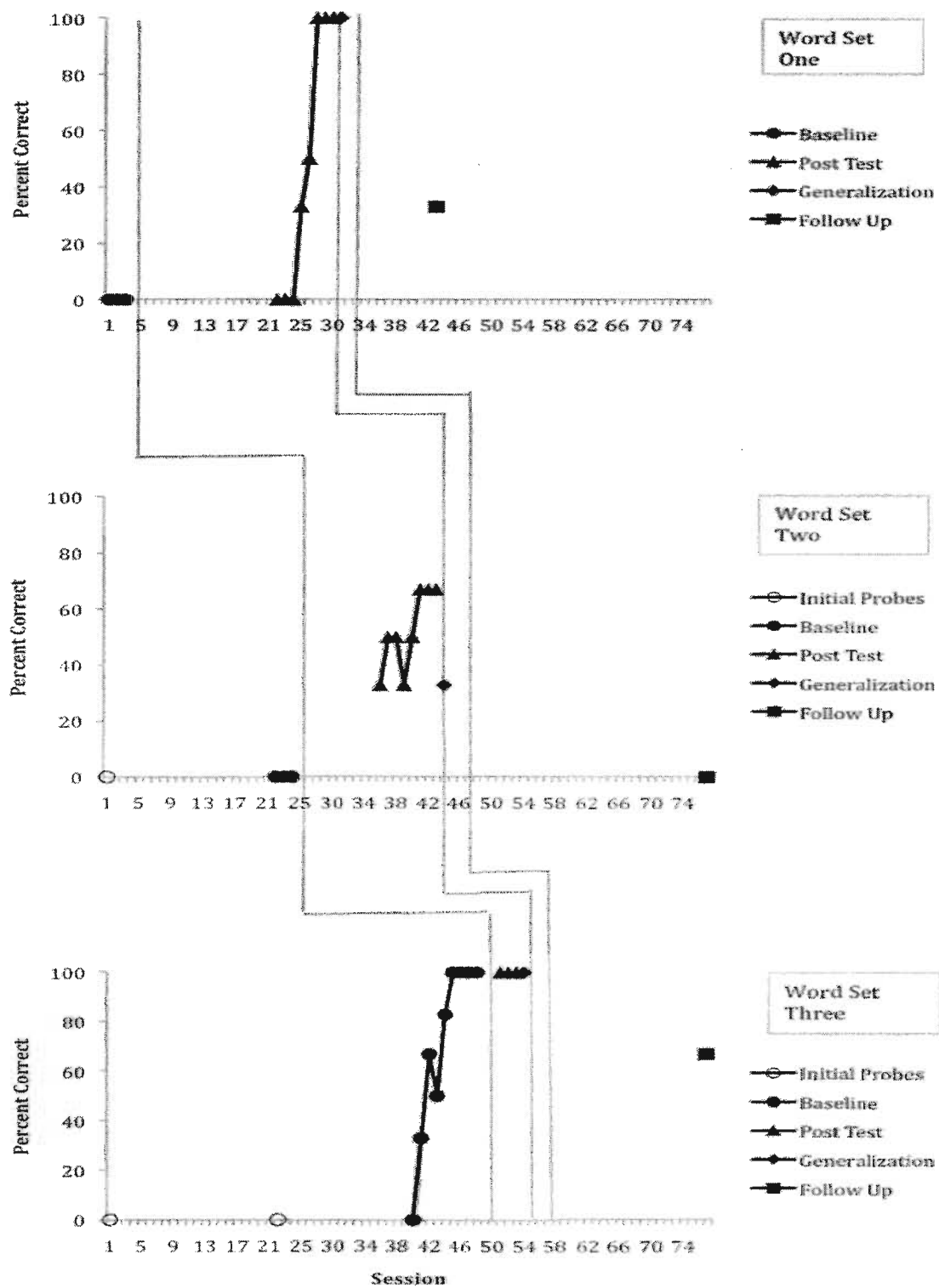


Figure 12. Participant 1: Oral naming of written word across all three word sets (CD)

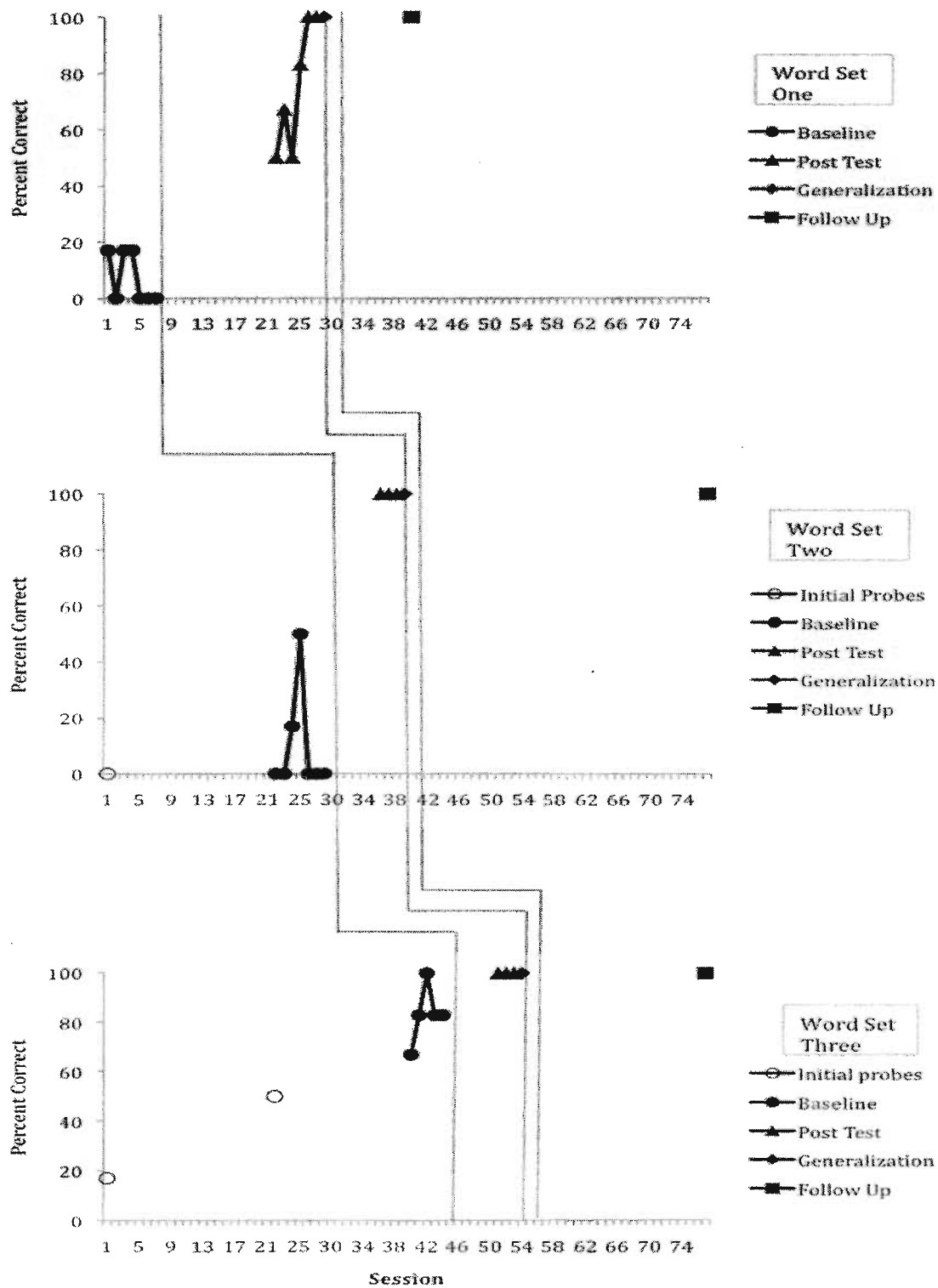


Figure 13. Participant 1: Matching picture to written word across all three word sets (BC)

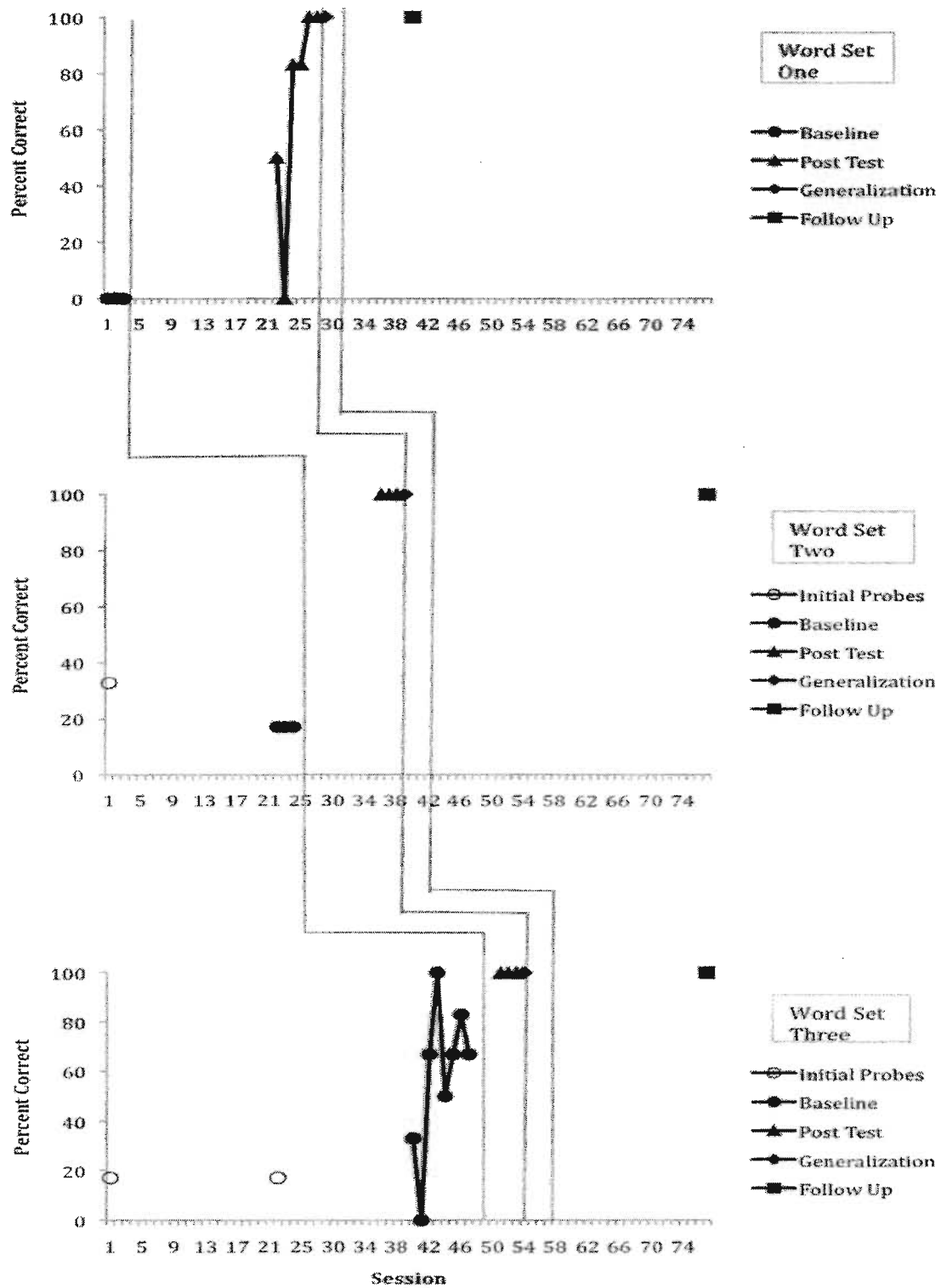


Figure 14. Participant 1: Matching written word to picture across all three word sets (CB)

Following baseline, James's Word Set One entered into the intervention phase. Figure 9 shows that a total of 11 intervention sessions, (10 trials per word) or a total of 330 trials were completed for his performance to meet the mastery criterion. For James's first three training sessions gestural prompts were used to prompt the correct response. However an increase in errors was reported as therapists attempted to fade the prompts. Prompting was therefore switched to physical (hand over hand) prompts, and the first trial was scored as independent the following session. The intervention phase continued until all three words met mastery.

James's original post test results (i.e. using the original test order) were gathered across three data points, and a total of 180 trials. Reflexivity results have already been discussed and can be found in Table 3. Results from the other relations indicated that James's accuracy remained high for matching dictated names to the pictures (BA) at 100%, as shown in Figure 10, and the skill which was taught, matching written words to dictated names (AC), as shown in Figure 9, was maintained with performance of 100% across all three data points. Accuracy on the remaining relations was not indicative of equivalence (or complete equivalence), based either on highly variable performance or accuracy at 0%. More specifically, for the oral naming of the written words (CD), as shown in Figure 12, James failed to name the word on each of the six trials presented to him. For matching the written words to pictures (CB) and the pictures to written words (BC) (Figure 14 and 13 respectively) performance was variable, with accuracy ranging from 0 to 83% for matching words to pictures and from 50% to 67% for matching pictures to words. Similarly, the symmetry test involving the dictated names and written words (CA), as shown in Figure 11 was also variable with accuracy ranging from 44% to 89%. Given these mixed results, James's post tests were repeated with the revised testing order previously described. Post tests were only repeated for relations for which accuracy was not 100% during the initial post tests. For part 1, the combined test for symmetry and

transitivity, James's demonstrated 100% accuracy on the review tests (AB and AC), conducted prior to the combined test (Note, the results of the AB test are not presented on the graphs as these tests were repeated as review for BC and CB tasks, this is also the case for all review relations that follow). The upper graphs of Figure 14 shows James's accuracy on matching written words to pictures (CB), and shows that following the review of AB and AC tasks James achieved criterion on this test with scores of 83%, 100% and 100% across the three sessions. Similarly, for matching pictures to written words (BC) as shown in Figure 13, accuracy was also 83%, 100%, and 100%. For Part 2 (Figure 12), the test for oral naming of the written words, James's demonstrated 100% accuracy on the reviewed relations (BD and BC) tested prior to the testing of oral naming of written words. James's accuracy for the oral naming of the written words (CD) emerged gradually following the review with scores of 33%, 50%, 100%, 100%, and 100%. For the symmetry test (CA) tested in Part 3 (Figure 11), the reviewed relation (AC) was maintained at 100% across all sessions. For CA task, James' initial accuracy was 67%, however, performance gradually increased to 100% across three sessions. With the revised order of the post test James passed all equivalence tests in a total of 297 trials across 8 days. With both the initial and revised post tests, testing took place across 14 days and 477 trials.

Following the post test phase, a generalization session was completed in which James repeated the tests for a number of relations using novel stimuli that were not presented during baseline, intervention or post test phases. Generalization sessions were completed for the oral naming of the written word (CD), matching the picture to the written word (BC), matching the written word to the picture (CB), and matching the written word to the dictated name (AC) (Figures 12, 13, 14, and 9). James's achieved 100% across all generalization probes, indicating novel stimuli included in the classes.

When Word Set Two completed post test and Word Set Three entered baseline, a follow-up probe was conducted for Word Set One in order to determine whether or not emergent relations were maintained. A follow up probe session was conducted for all relations. The upper graphs of Figures 9,13, and14 show that James's performance in matching written words to dictated names (AC), written words to pictures (CB), and matching pictures to written words (BC), respectively, were maintained at follow up. Similarly, James's performance for the BA and CA tasks (Figures 10 and 11) were well maintained at follow up. Finally, Figure 12 shows that oral naming of written word (CD) was maintained for only the word ("baby").

Word Set Two. The words two, bike, and five were denoted as Word Set Two. James's performance on Word Set Two for each phase of the study (including the single probe session given when Word Set One was in baseline) will be discussed next.

A single probe session was completed for Word Set Two when Word Set One was in baseline. Figure 9 shows that James's performance on matching written words to dictated names (AC), the skill that was taught, was 0% on the initial probe. For matching dictated names to pictures (BA), and matching dictated names to written words (CA), James's accuracy at the initial probes was 89% and 22% respectively (Figures 10 and 11). Figures 12-14 show the remaining emergent relations, oral naming of the written words (CD), matching pictures to the written words (BC), and matching written words to pictures (CB). Accuracy for both the CD and BC tasks was 0% at the initial probe, whereas accuracy for the CB task was 33%.

In order to obtain a stable baseline for James's second word set across all the relations, 366 test trials were completed. Figure 9 (middle graph) shows that James's performance for matching written words to dictated names (AC) was initially variable but eventually stabilized at 0%. The middle graph of Figure 10 shows that James's accuracy on the matching of dictated names to the pictures (BA) was nearly perfect. The middle

graph of Figure 11 shows that James's accuracy for matching dictated names to written words (CA) was quite variable but stabilized at 11%. James's performance in oral naming of written words (CD) (Figure 12) was 0% across three sessions and in matching written words to the pictures (CB) (Figure 13) was also low. James's performance in matching the picture to the word (BC) (Figure 14) was initially variable but eventually stabilized at 0%.

Following baseline, the intervention phase started. The middle graph of Figure 9 shows that James mastered Word Set Two in only seven sessions. A total of 210 trials were completed to reach mastery criterion. Where needed, physical prompts were used to prompt correct responding.

James's post tests for Word Set Two were run in the revised order already discussed. For part 1, the combined test for symmetry and transitivity, James's demonstrated 100% accuracy on the review tests (AB and AC) tested prior to the combined test. Figure 14 shows that, following the review, James achieved criterion immediately for matching written words to pictures (CB) with scores of 100% across the three sessions. Similarly, perfect accuracy scores were demonstrated (Figure 13) for matching pictures to written words (BC). The middle graph of Figure 12 shows the results from part 2 which entailed a review of picture naming (BD) and matching pictures to written words (BC), followed by testing the oral naming of written words (CD). James's accuracy on both reviewed relations was 100%. James's accuracy for the oral naming of written words (CD) ranged from 33% up to 67% and eventually stabilized at 67%. James correctly named "two" across the last three sessions, he correctly labeled "bike" across two of the three sessions. "Five" however was only correctly labeled for one of the last three sessions. For part 3, symmetry of dictated names and written words, James's performance on AC, the relation being reviewed (Figure 9) was maintained at 100% across all sessions. Following the review test, James's accuracy for CA was near

perfect (Figure 11),. Similarly, for part 4 (symmetry of dictated names and pictures) James's demonstrated perfect accuracy on the relation being reviewed (AB) and Figure 10 shows perfect performance for the symmetry of this relation (BA) as well. Finally, part 5 tested the reflexivity (Table 3) by testing matching names to names (AA), pictures to pictures (BB), and words to words (CC), as was already discussed. James passed all equivalence tests, and named some of the written words, in a total of 375 trials across 4 days.

Following the post test phase, a generalization phase was completed. The middle graphs of Figures 9,12, 13, and 14 show that James's achieved 100% across all generalization probes for Word Set Two, except for oral naming (Figure 12) indicating the majority of emergent relations were also generalized to novel stimuli. For oral naming (CD) his performance for "two", the word that was most consistently correct during post test, was generalized to the novel stimuli.

When James completed the post test for Word Set Three, a follow-up probe was to be conducted for Word Set Two. However, James was discharged from the IBI program immediately following the post test for Word Set Three. Consequently, follow up probes for Word Sets Two and Three were conducted simultaneously in James's home approximately 4 weeks following the post test for Word Set Three. The middle graphs of Figures 13, 14, and 9 show that James's performance for matching pictures to written words (BC), written words to pictures (CB), and matching written words to dictated names (AC), for Word Set Two, were maintained at follow up. James's performance for the BA and CA tasks (Figures 10 and 11) also demonstrates performance maintained at follow up, although accuracy was not perfect. Finally, Figure 12 shows that James's oral naming of written words (CD) for Word Set Two was not demonstrated at the follow-up; he scored 0%. It is important to note that for Word Set Two James's performance was variable during post test for the CD task as well.

Word Set Three. The words selected cat, wolf, and book were denoted as Word Set Three. James's performance on Word Set Three for each phase of the study will be discussed next.

A single probe session was completed for Word Set Three when Word Set One was in baseline and a second probe was completed when Word Set Two was in baseline. Figure 9 shows that James's performance for matching written words to dictated names (AC) was 0% at both the initial and second probe. James's accuracy levels for the single probe sessions for matching dictated names to pictures (BA) and matching dictated names to written words (CA) show at 89% (Figure 10) and 11% (Figure 11) respectively. Figure 12 shows that James's performance for oral naming of written words (CD) was 0% at the initial probes. Matching pictures to written words (BC), as depicted in Figure 13, was 17% for the first probe session and 50% for the second probe session and matching written words to pictures (CB) was 17% (Figure 14) for both probe sessions.

In order to obtain a stable baseline for James's third word set across all the relations, 204 test trials were completed. Figure 9 shows James's performance on the skill that was taught, matching written words to dictated names (AC), and indicate that James's performance ranged in accuracy from 0% to 67% and stabilized at 67%. The bottom graphs of Figure 10 shows that performance for matching dictated names to pictures (BA) was nearly perfect. Similarly, performance for matching dictate names to written words (CA) was relatively stable with Figure 11 depicting accuracy levels between 78% and 100%. Figure 12 shows that James's performance in oral naming of written words (CD) was initially 0%, however his level of accuracy increased across three sessions. Testing continued in order to achieve a stable baseline however James's accuracy eventually increased to 100% across three sessions. James's performance in matching pictures to words (BC) ranged from 67% to 100% (Figure 13) but eventually stabilized at 83%. Similarly, for matching written words to pictures (CB) performance

was also highly variable ranging from 0% to 100% (Figure 14). Baseline ended when James's performance stabilized at 67%.

Following baseline, James's Word Set Three entered into the intervention phase. Figure 9 shows that James mastered the intervention phase for Word Set Three in only 4 sessions (a total of 120 trials). Where needed, physical prompts were used to prompt correct responding.

Following the intervention phase, James's post test for Word Set Three began. For Part 1 (combined test for symmetry and transitivity) James demonstrated 100% accuracy for the review skills, AB and AC (Figure 9), tested prior to the combined test. The bottom graphs of Figures 14 and 13 show the results from matching words to pictures (CB) and matching pictures to words (BC), which was assessed following the review. James's performance for both relations was 100%. Perfect accuracy was also found for the reviewed skills for Part 2 (BD and BC). Figure 12 shows James's accuracy for oral naming of written words (CD), which was 100% across all sessions. For part 3, James's performance on the relation being reviewed (AC) was maintained at across all sessions, with the exception of 1 error made during which James touched the wrong word when asked to touch "cat". For the symmetrical relation, CA (Figure 11), however, James demonstrated 100% across three sessions. Similarly, for part 4 James's performance on the relation being reviewed (AB) was again perfect as was the symmetry of this relation (BA) depicted in Figure 10. Finally, part 5 tested the reflexivity by testing matching names to names (AA), pictures to pictures (BB), and words to words (CC) (Table 3), as was already discussed. James passed all equivalence tests in a total of 270 trials across 2 days.

Following the post tests, a generalization session was completed. Figures 9, 12, 13, and 14 show that James's achieved 100% across all generalization probes for Word

Set Three, indicating the majority of his emergent performances were also generalized to novel stimuli.

As was already discussed, follow up probes for Word Set Three were conducted in James's home approximately 4 weeks following the post test for Word Set Three. Figures 13, 14, and 9 show that James's performance for matching pictures to written words (BC), matching written words to pictures (CB) and matching written words to dictated names (AC) were maintained at follow up. James's performance for the BA and CA tasks (Figures 10 and 11) also demonstrates performance maintained at follow up, although this maintenance imperfect. Finally, Figure 12 shows that oral naming of written words (CD) for Word Set Three was maintained at follow up for two of the three words ("cat" and "wolf" were maintained, but that "book" was not).

In order to complete all phases of the study, for all three word sets, James required a total of 72 days. Training and testing for Word Set One was completed in 38 days, Word Set Two in 18 days and Word Set Three in 11 days.

Participant 2

Reflexivity. Table 4 shows the initial baseline, post test and follow up results for the identity matching tasks (AA, BB and CC). Billy's performance at the initial baseline probes was perfect, across all three word sets, for both matching names to names (AA) and matching pictures to pictures (BB). For matching words to words (CC) Billy's accuracy scores for the initial baseline probes was 50% for the first word set, and 33% for the second and third word sets. For post test and follow up phases, Billy performed perfectly on all reflexivity tests. Matching words to words (CC) emerged following the acquisition of matching written words to dictated names.

Table 4:

Participant 2: Reflexivity scores for all three word sets

Word Set	<u>Name to Name (AA)</u>			<u>Picture to Picture (BB)</u>			<u>Word to Word (CC)</u>		
	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up
1	100%	100%	100%	100%	100%	100%	50%	100%	100%
2	100%	n/a	100%	100%	n/a	100%	33%	n/a	100%
3	100%	100%	100%	100%	100%	100%	33%	100%	100%

Word Set One. From the nine words selected for Billy, ant, cup, and dog were denoted as Word Set One. Recall that Word Set One entered baseline and at the same time a single probe session was completed for Word Sets Two and Three. Results of Billy's Word Set One for each phase of the study will be discussed next, and the results of the single probe sessions for Word Sets Two and Three will be discussed later when presenting the results for these word sets.

In order to obtain a stable baseline for Billy across all the relations, 351 test trials were completed. The upper graph of Figure 15 shows that matching written words to dictated names (AC), the skill that was taught, stabilized at 0% in sessions three through five. For matching dictated names to pictures (BA) Billy's score was initially 56% (Figure 16), however, despite the absence of reinforcement, increased over 9 sessions and was stable at 100% across the last three sessions. For matching dictate names to written words (CA), Billy's performance was for the first six sessions of baseline was on a steady upward trend (Figure 17). An error analysis was completed and following the modifications already described, Billy's performance decreased (44%, 22%, 11%). The results for oral naming of written words (CD), matching pictures to written words (BC) and matching written words to pictures (CB) are shown in the upper graphs in Figure 18, 19, and 20 respectively. Billy's scores on all of these tasks was 0% across all sessions.

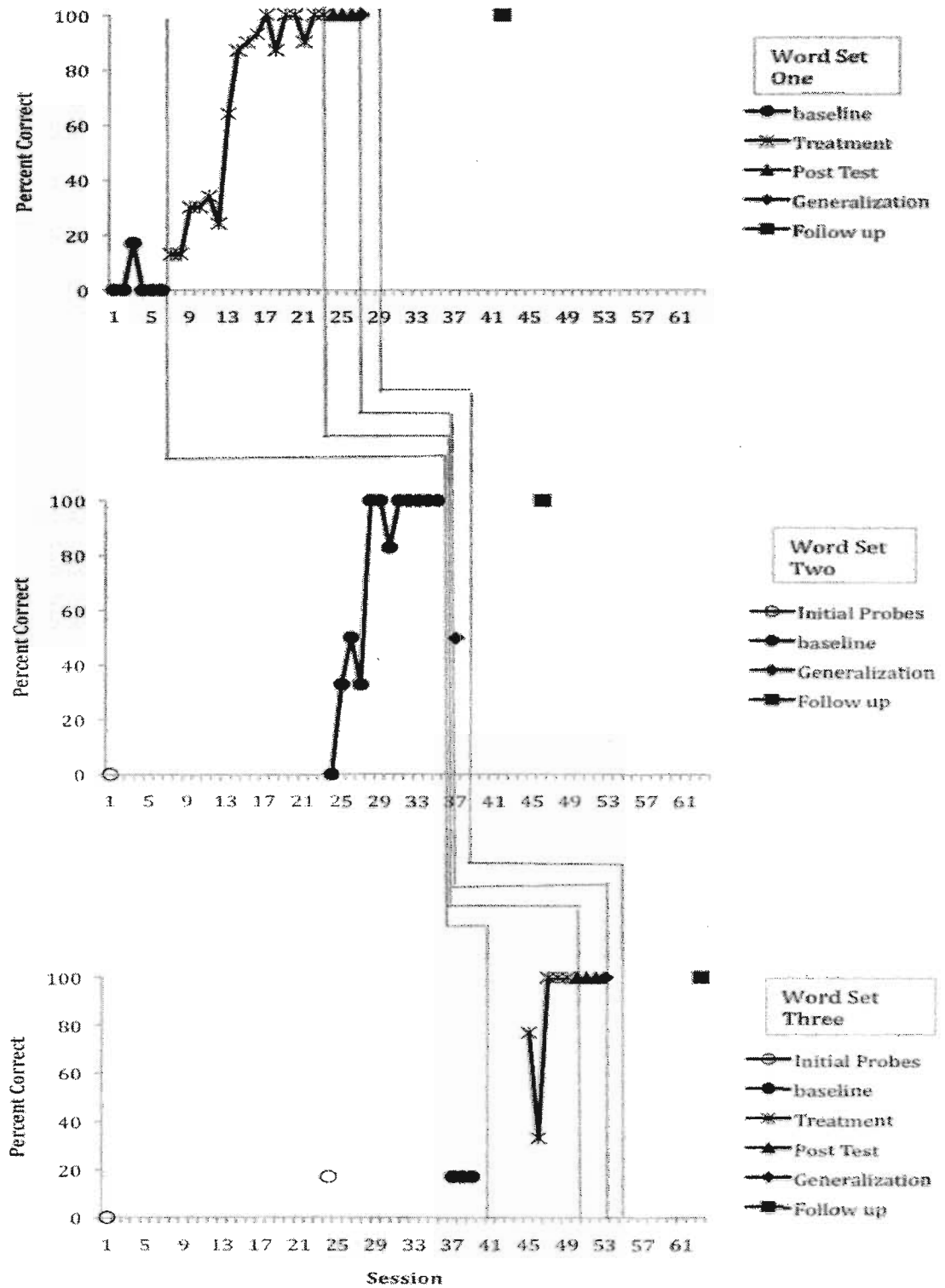


Figure 15: Participant 2: Matching written word to dictated name across all three word sets (AC-Taught relation)

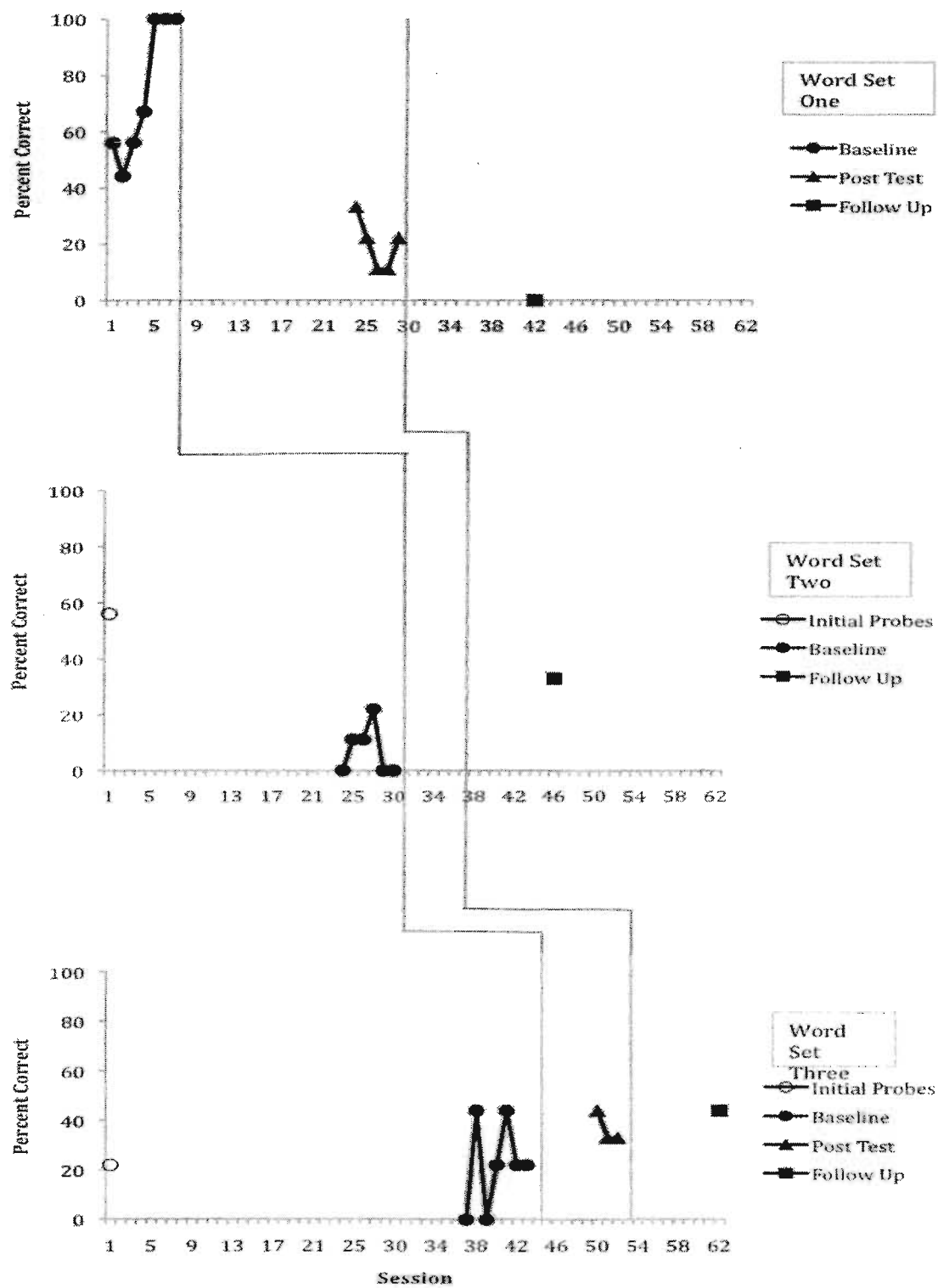


Figure 16: Participant 2: Matching dictated name to picture across all three word sets

(BA)

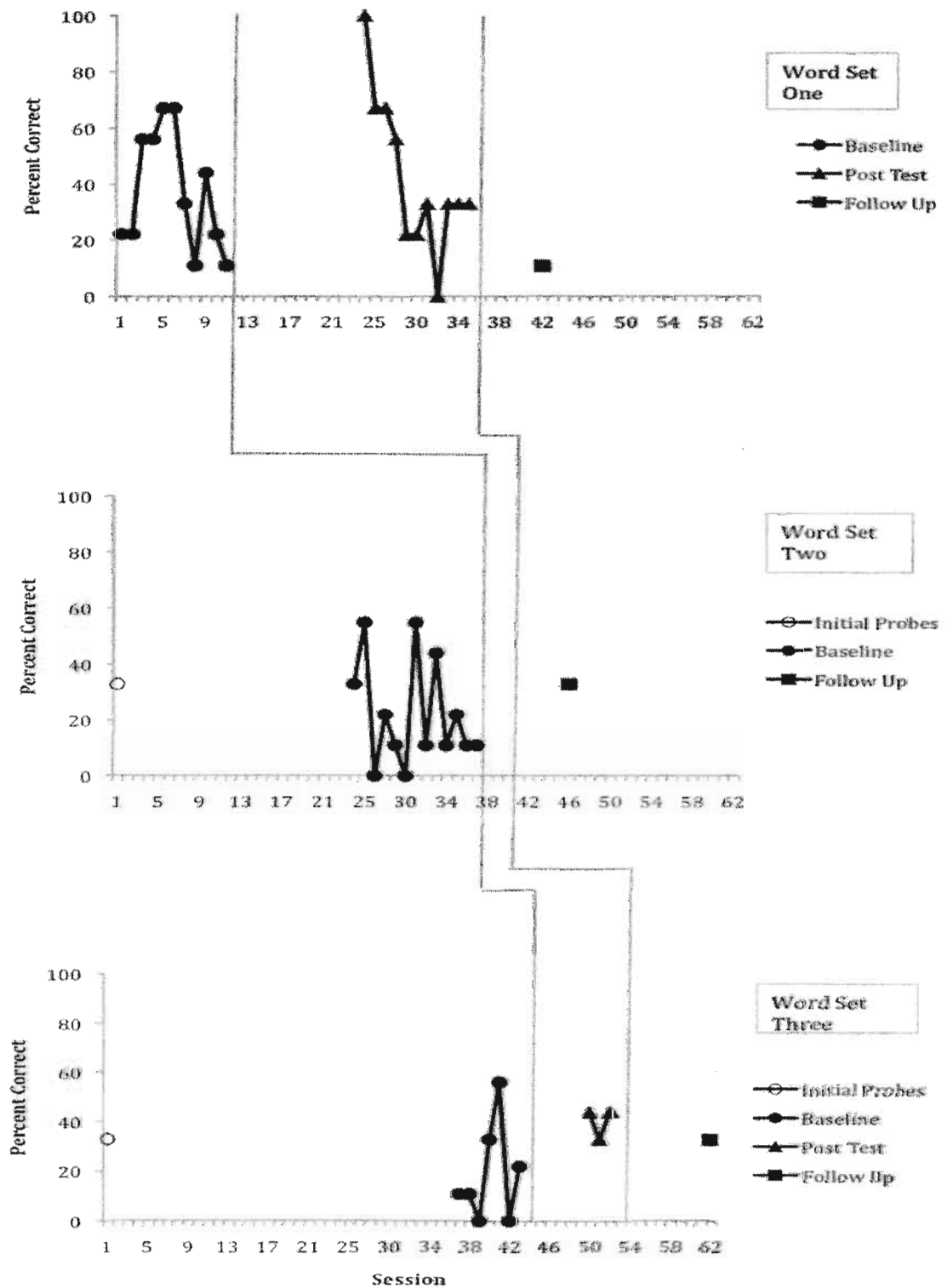


Figure 17: Participant 2: Matching dictated name to written word across all three word sets (CA)

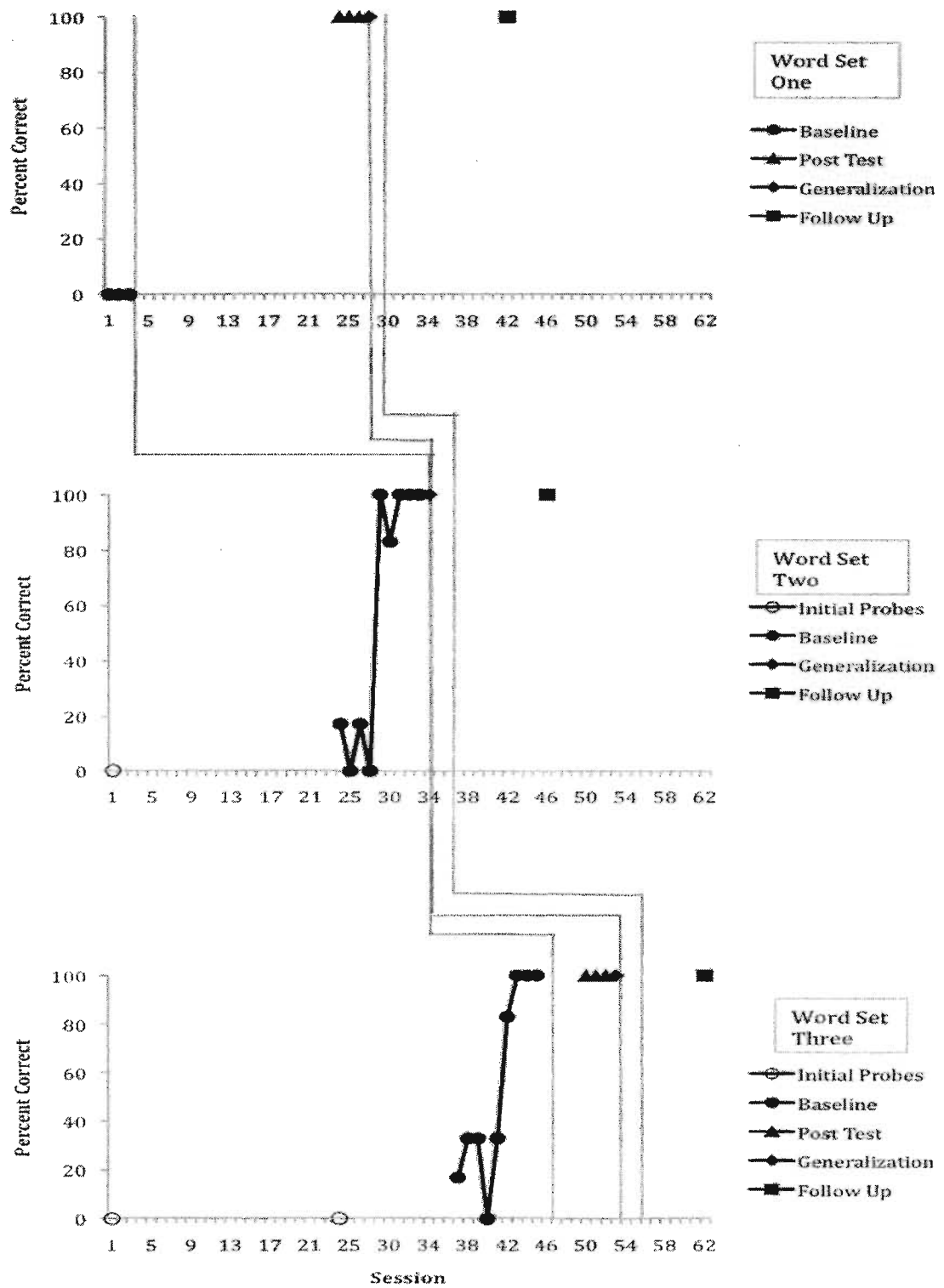


Figure 18: Participant 2: Oral naming of written word across all three word sets (CD)

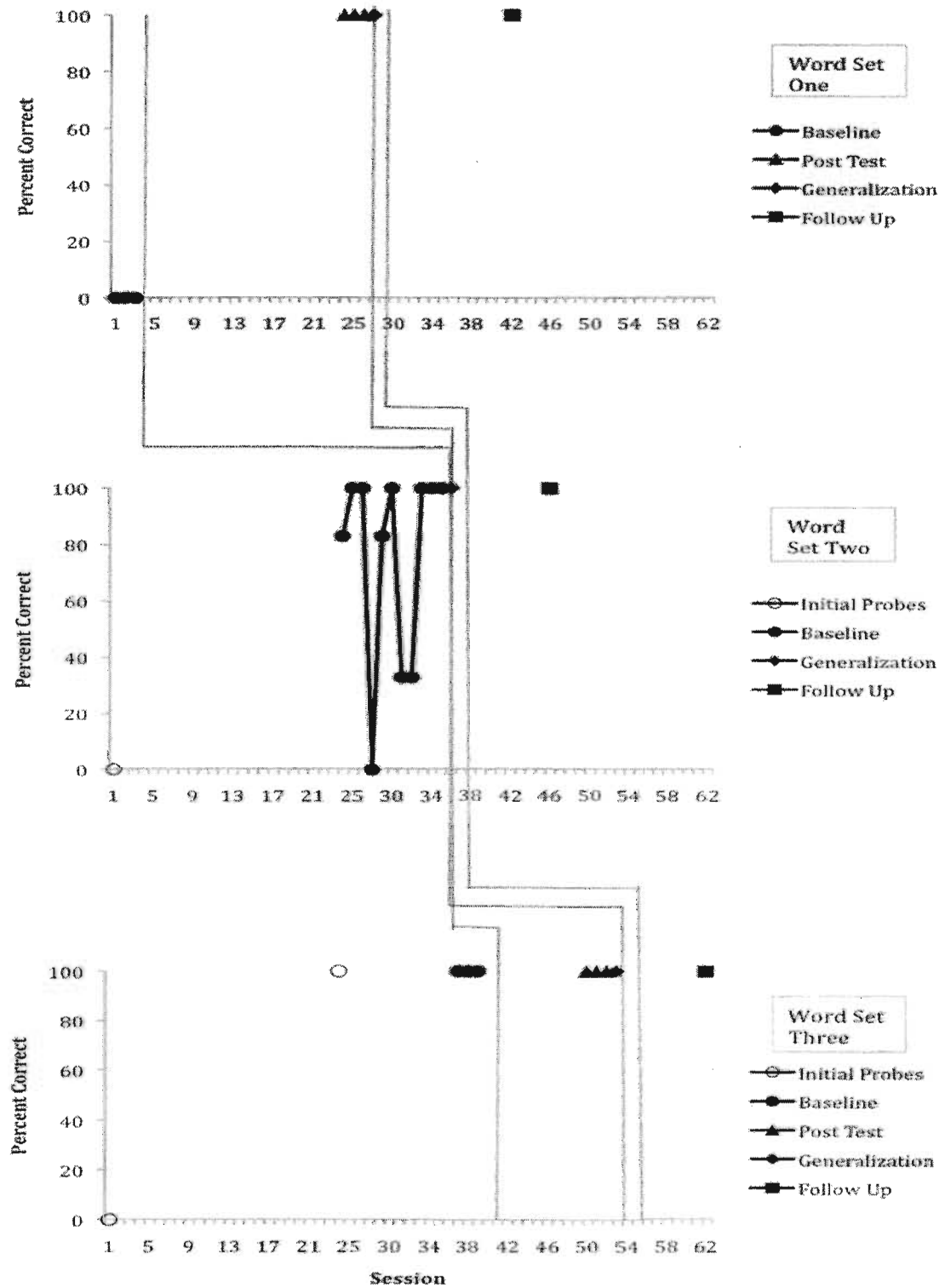


Figure 19: Participant 2: Matching picture to written word across all three word sets (BC)

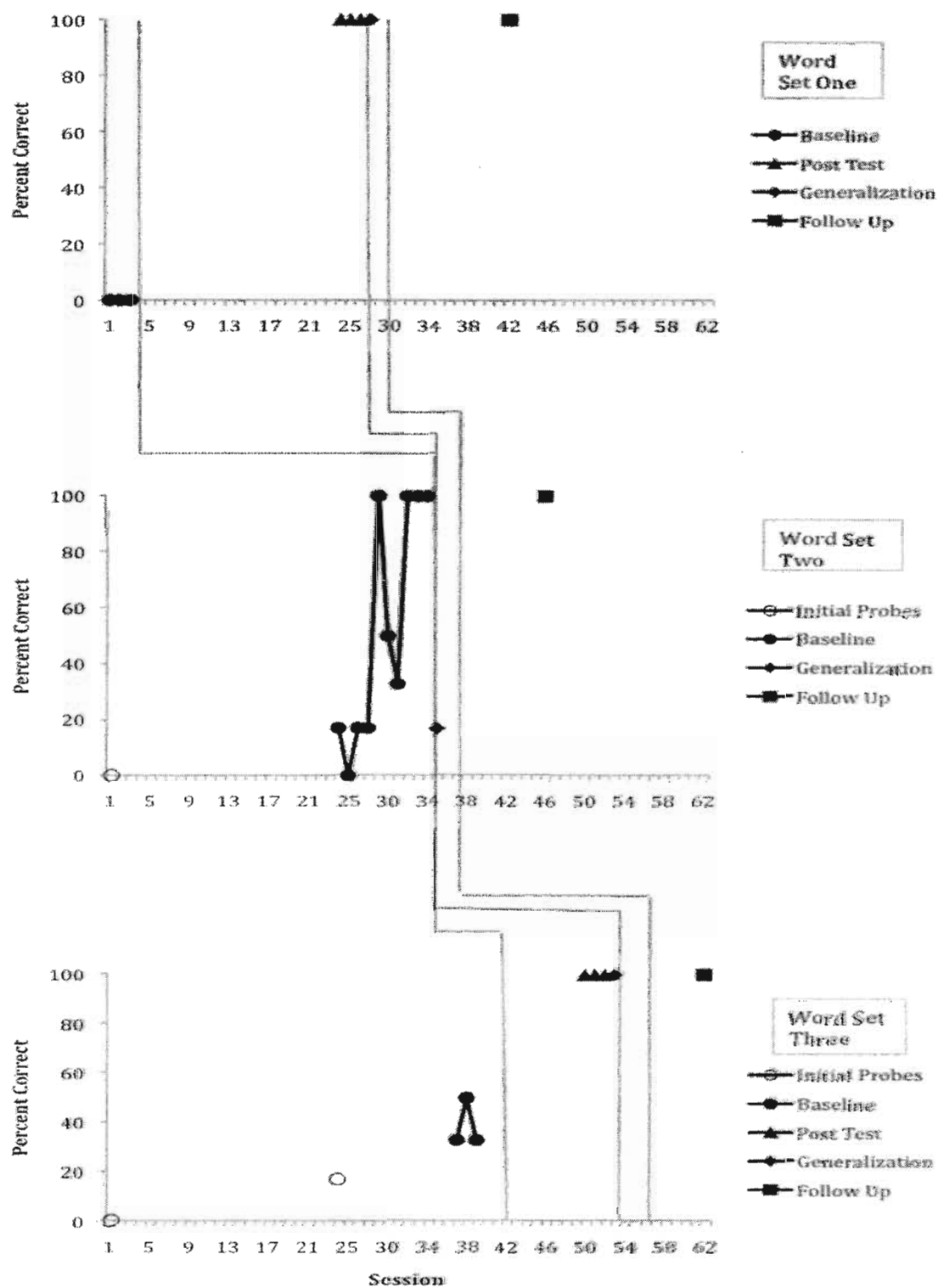


Figure 20: Participant 2: Matching written word to picture across all three word sets (CB)

Following baseline, the intervention phase began. Figure 15 shows that a total of 17 intervention sessions, (a total of 510 trials) were completed prior to Billy's post test for Word Set One (however, in analyzing the data it was actually found that mastery criterion was met at day 14). A combination of physical and gestural prompts was used throughout the intervention phase for Billy. Two words reached mastery at a faster rate than the third word. Therefore, all three words were continued until the third word reached the mastery criterion.

Following the intervention phase, Billy's post test phase was run. For part 1, the combined test for symmetry and transitivity, Billy's demonstrated 100% accuracy on the review tests (AB and AC) conducted prior to the combined tests. Billy also demonstrated 100% accuracy for (CB), the matching written words to pictures (Figure 20) and for (BC), matching pictures to written words (Figure 19). Part 2 entailed testing the oral naming of written words (CD). Billy's had perfect performance on the reviewed relations (BD and BC). Figure 18 shows that, following the review, Billy's performance for oral naming of written words (CD) was also 100%. For the symmetry of the written words and dictated names, tested in Part 3, Billy's accuracy on the relation being reviewed (AC) was 100%. For the CA task (Figure 17), Billy's first session of post test was indicative of symmetry with accuracy of 100%. On subsequent sessions however his accuracy declined rapidly before stabilizing at 33%. Observation indicated that Billy was matching the sample stimulus to the middle comparison stimulus on every trial. Part 3 was tested across 99 trials before discontinuing testing and moving on to Part 4. For Part 4, Billy's demonstrated perfect performance for the review skill (AB). Testing then proceeded with testing matching dictated names to pictures (BA), which had relatively low performance ranging from 11% to 33% (Figure 16). Observation of this testing indicated that Billy was matching the sample stimulus to the middle comparison stimulus on almost every trial, just as he did with the CA task. Lastly, Part 5 tested the reflexive relations (AA, BB,

and CC), which have already been discussed (Table 4). Billy completed the post tests for Word Set One in a total of 384 trials across 11 days.

Following the post test phase, a generalization session was completed. Figures 18, 19, 20, and 15 show that Billy achieved 100% across all generalization probes for Word Set One, indicating that his emergent performances were also generalized to novel stimuli.

When Billy completed the post test for Word Set Two and Word Set Three entered baseline, a follow-up probe was conducted for Word Set One. Figures 19, 20, 18, and 15 show that matching pictures to written words (BC), matching written words to pictures (CB), oral naming of written words (CD), and matching written words to dictated names (AC) were maintained at follow up. Billy's performance for the BA and CA tasks (Figures 16 and 17) show similar results as to what was found in post test, with accuracy levels of 11% and 0% respectively.

Word Set Two. The words "Fish", "Owl" and "Door" were denoted as Word Set Two. Results of Billy's Word Set Two for each phase of the study (including the single probe session when Word Set One was in baseline) will be discussed next.

Figure 15 shows that matching written words to dictated names (AC), the skill that was taught, was 0% at the initial probe. Matching dictated names to pictures (BA) and matching dictated names to written words (CA) were, 56% (Figure 16) and 33% (Figure 17) respectively at the initial probe. Figures 18, 19, and 20 show that the remaining tasks, oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB) were all at 0% on these probes.

In order to obtain a stable baseline for Billy's second word set across all the relations, 159 test trials were completed. Interestingly, Billy's performance for matching written words to dictated names (AC), the skill to be taught, stabilized at 100% in baseline, without any teaching and despite the skill not being in his repertoire during the

initial probes (Figure 15). Figure 16 shows that Billy's performance on matching dictated names to pictures (BA) was variable at or below chance levels with accuracy scores ranging from 0-22%. Similarly, Figure 17 shows that for matching dictate names to written words (CA), performance varied between 0% and 55% with a downward trend for the last three sessions. The middle graphs of figures 18, 19, and 20 show the results of oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB), which were all initially variable but stabilized at 100% across the last three sessions.

Following baseline, Billy's Word Set Two was to enter into the intervention phase. However, given that Billy's baseline responding for matching written words to dictated names (AC) was already at mastery and all relations with the exception of matching dictated names to pictures and dictated names to written words were also at mastery, the intervention and post test phases for Word Set Two were not completed.

Following the baseline therefore, a generalization session was completed. Figures 18 and 19 show that Billy's achieved 100% for oral naming of written words (CD) and for matching pictures to written words (BC) across all generalization probes, indicating his emergent performances were also generalized to novel stimuli. For matching written words to pictures (CB) and matching written words to dictated names (AC) however, performance did not generalize, with scores of 16.7% (Figure 20) and 50% (Figure 15) respectively.

When Billy completed the post test for Word Set Three, a follow-up probe was conducted for Word Set Two. Figures 15, 18, 19, and 20 show that Billy's performance for matching written words to dictated names (AC), oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB) were maintained at follow up. Billy's performance for the BA and CA tasks (Figures 16

and 17) demonstrates similar poor performance to that demonstrated during the post test phase.

Word Set Three. The words “milk”, “lion” and “star” were denoted as Word Set Three. Results of Billy’s Word Set Three for each phase of the study will be discussed next.

Figure 15 shows that Billy’s accuracy score for matching written words to dictated names (AC) was 0% at the initial probe and 17% at the second probe. Figures 16 and 17 show that Billy’s accuracy for matching dictated names to pictures (BA) and matching dictated names to written words (CA) at the initial probe sessions were 22% and 55% respectively. For the second probe sessions, results for both of these relations were at 11%. The results of oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB), show that Billy’s accuracy scores were at 0% on these relations for the initial probe (Figures 18, 19, and 20). At the second probe however, oral naming of written words (CD) was still at 0%, and matching written words to pictures was at 17%, but matching pictures to written words (BC) was at 100%.

In order to obtain a stable baseline for Billy across all the relations, 387 test trials were completed. Billy’s performance on the skill to be taught, matching written words to dictated names (AC), was 17% (Figure 15). The bottom graph of Figure 16 shows that Billy’s performance on matching dictated names to pictures (BA) was variable between 0% and 67%, and was discontinued following a steady downward trend. The bottom graph of Figure 17 shows that Billy’s accuracy on matching dictated names to written words (CA) was varied between 0% and 55%. Error analysis indicated that Billy’s was matching the sample stimulus to two consistent locations for the comparison stimuli. The therapists were therefore instructed to avoid those positions and a downward trend was observed across three sessions. Figure 18 shows that Billy’s performance in oral naming

of written words (CD) varied between 17% and 33%. Baseline continued in order to get stability and accuracy increased to 100%. Figure 19 shows that accuracy for matching pictures to written words (BC) was stable at 100%. Lastly, Figure 20 shows that Billy's accuracy levels for matching written words to pictures (CB) ranged between 33% and 50%.

Despite Billy's high performance on many of the baseline relations, his performance on the relation to be taught, matching written words to dictated names (AC) was only 16.7%. Consequently, following baseline, Billy's Word Set Three entered into the intervention phase. Figure 15 shows that a total of five intervention sessions, (a total of 150 trials) were completed in order to reach mastery criterion. Physical prompts were used as needed throughout the intervention phase for Billy.

Following the intervention phase Billy's Word Set Three began the post test phase. For Part 1 (combined test for symmetry and transitivity), Billy demonstrated perfect performance on the review tests (AB and AC). Figure 20 shows that for matching written words to pictures (CB), which followed, Billy achieved criterion immediately with scores of 100%. Similarly, Figure 19 shows perfect performance for matching pictures to written words (BC). For Part 2 (oral naming of written words) Billy demonstrated perfect performance on the reviewed relations (BD and BC). Figure 18 shows that Billy's then demonstrated 100% accuracy for the oral naming of written words (CD). Part 3 again reviewed the matching of written words to dictated names (AC), which was maintained at 100%. This was followed by testing the matching of dictated names to written words (CA). Figure 17 shows that for CA, Billy's performance was low with 44%, 33%, and 44% across sessions. For Part 4 Billy demonstrated 100% performance for the reviewed skill (AB). Testing then proceeded with testing matching dictated names to pictures (BA) with relatively low performance of 44%, 33%, and 33% (Figure 16). Lastly, Part 5 tested the reflexive relations (AA, BB, and CC), which were

already discussed (Table 4). Billy completed the post tests for Word Set Three in a total of 273 trials across 4 days.

Following the post tests, a generalization session was completed. Figures 18, 19, 20, and 15 show that Billy achieved 100% across all generalization probes for Word Set Three, indicating that his emergent performances were also generalized to novel stimuli.

Ten days following the completion of Billy's post test for Word Set Three, a follow-up probe was conducted for Word Set Three. Figures 15, 18, 19, and 20 show that Billy's performance for matching written words to dictated names (AC), oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB) were maintained at follow up. Billy's performance for the BA and CA tasks (Figures 16 and 17) demonstrates similar performance to those attained during the post test, with scores of 33% and 44% respectively.

In order to complete all phases of the study, for all three word sets, Billy required a total of 67 days. Teaching and testing for Word Set One was completed in 37 days, Word Set Two in 9 days, and Word Set Three in 19 days.

Participant 3

Reflexivity. Table 5 shows the initial baseline, post test and follow up results for the identity matching tasks (AA, BB, and CC). Henry's performance at the initial baseline probes, across all three word sets, for both matching names to names (AA) and matching pictures to pictures (BB) was perfect. For matching words to words (CC) Henry's performance at the initial baseline probes was 0% for the first word set, 17% for the second word set, and 33% for the third word set. For post test and follow up phases, Henry's reflexivity tests were all perfect.

Table 5:

Participant 3: Reflexivity scores for all three word sets

Word Set	<u>Name to Name (AA)</u>			<u>Picture to Picture (BB)</u>			<u>Word to Word (CC)</u>		
	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up
1	100%	100%	100%	100%	100%	100%	0%	100	100
2	100%	n/a	100%	100%	n/a	100%	17%	n/a	100
3	100%	n/a	100%	100%	n/a	100%	33%	n/a	100

Word Set One. The words bike, pig, and glue were denoted as Word Set One.

Results of Henry's word set one for each phase of the study will be discussed next, and the results of the single probe sessions for Word Sets Two and Three will be discussed later when presenting the results for these word sets.

In order to obtain a stable baseline for Henry across all the relations, 228 test trials were completed. Figure 21 shows that matching written words to dictated names (AC), the skill that was taught, was at 0% during baseline. For both matching dictated names to the pictures (BA) and matching dictate names to the written words (CA) Henry's performance was 33% across all sessions (Figure 22 and 23). Informal analysis indicated a clear positional preference, thus Henry got one response correct per word when the correct comparison stimulus was in that position. The upper graphs in Figure 24, 25 and, 26 show that oral naming of written words (CD), matching pictures to written words (BC) and matching written words to pictures (CB) were all at 0% across all sessions.

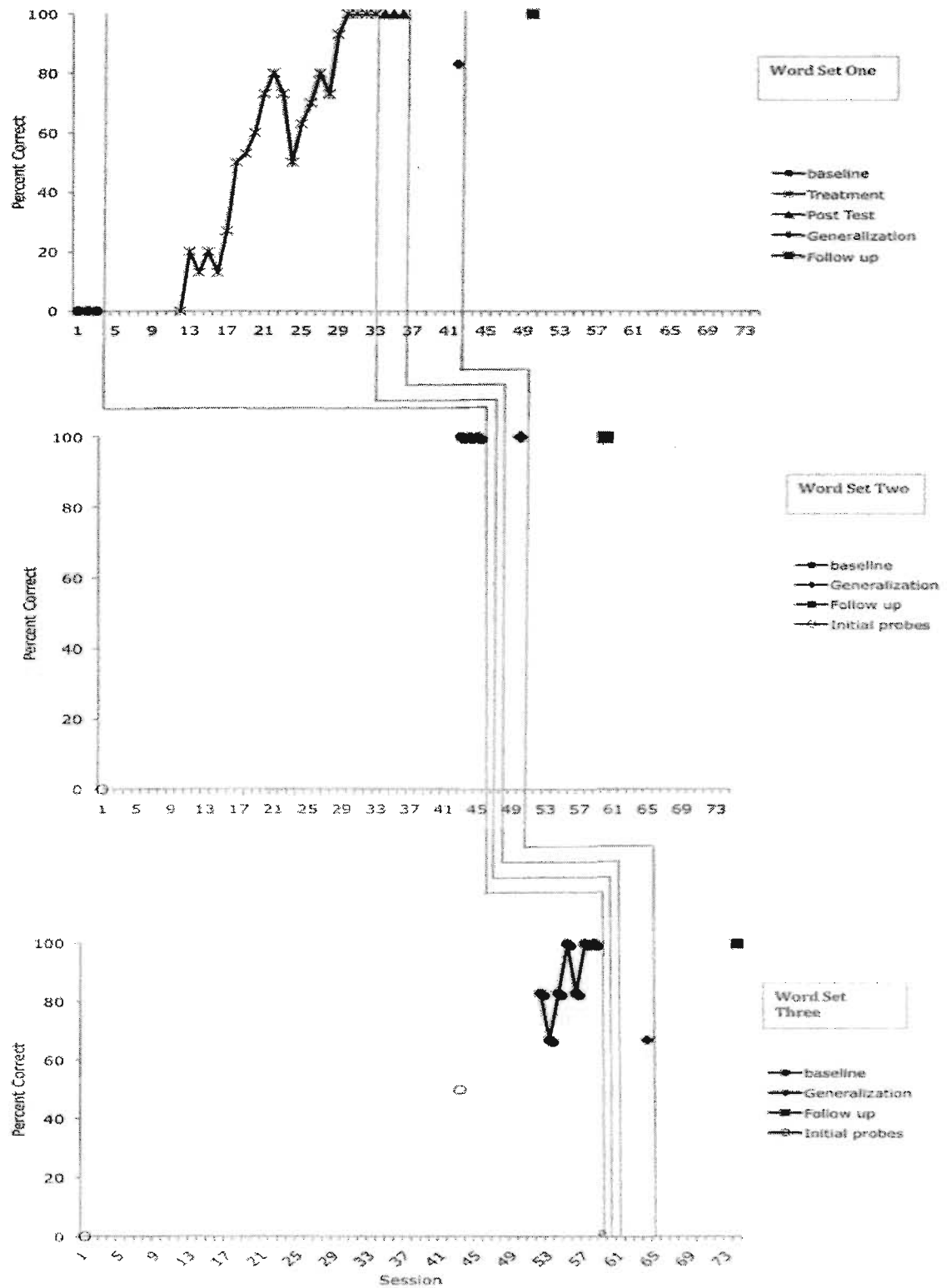


Figure 21: Participant 3: Matching written word to dictated name across all three word sets (AC-Taught relation)

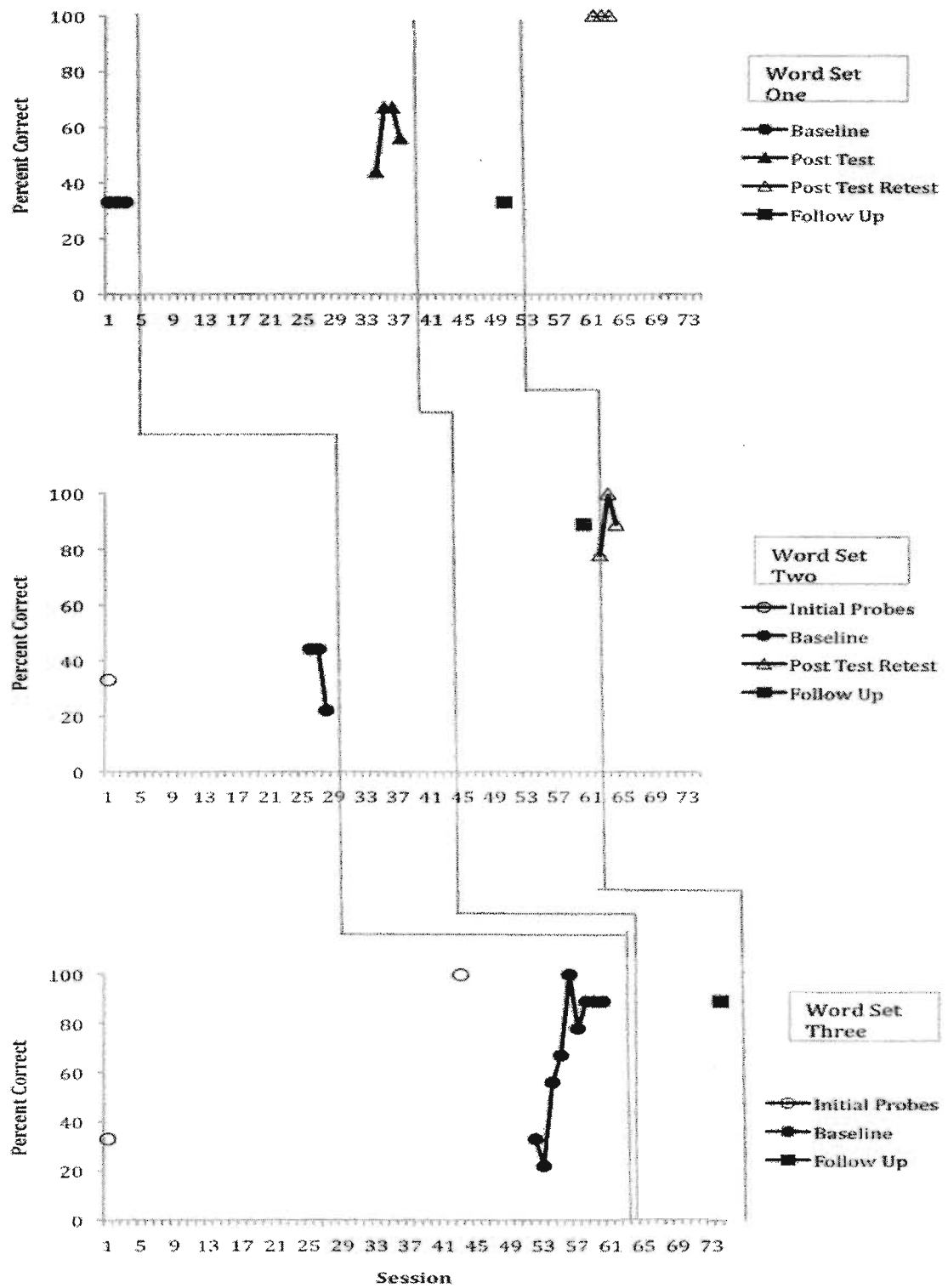


Figure 22: Participant 3: Matching dictated name to picture across all three word sets

(BA)

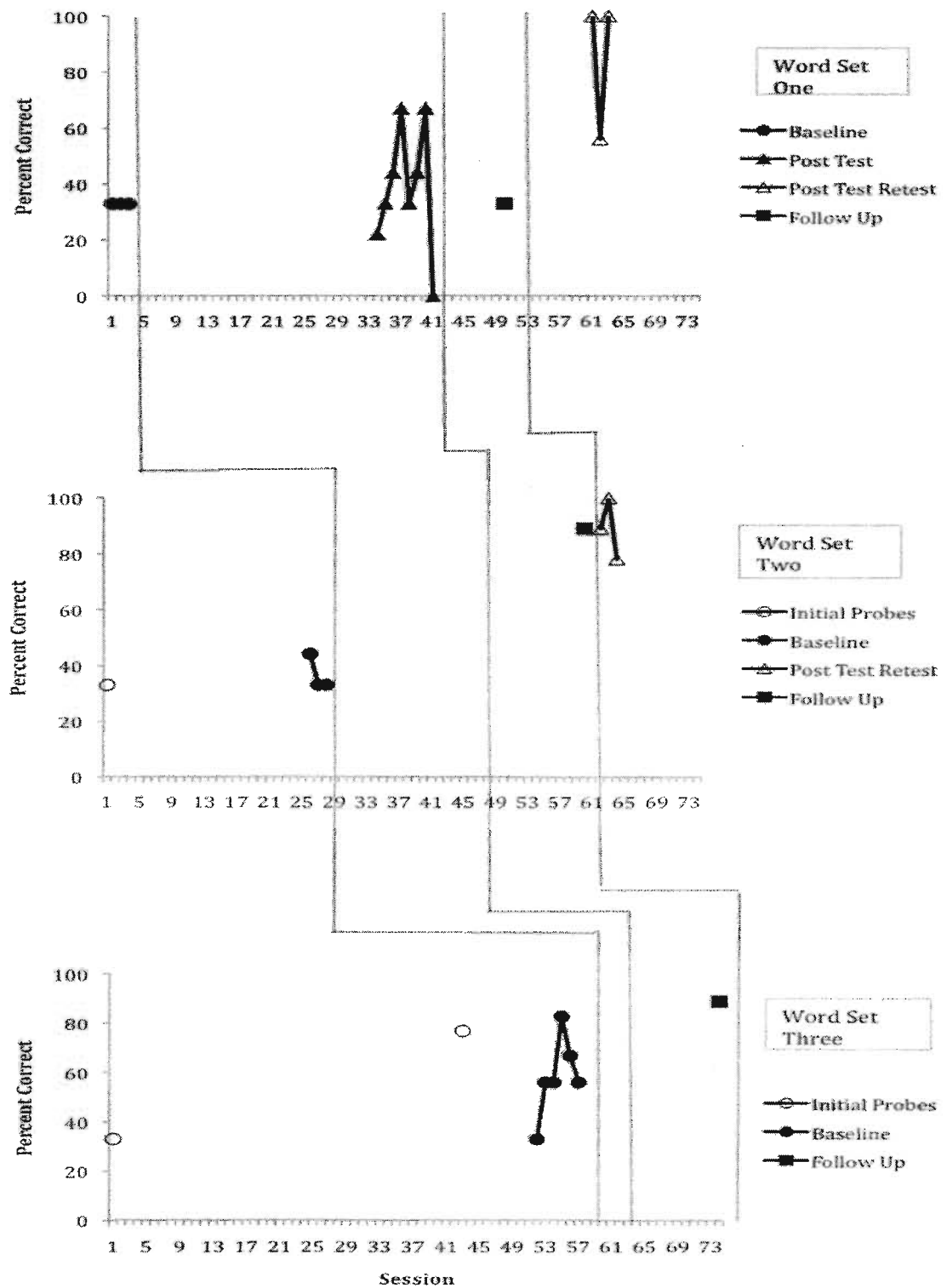


Figure 23: Participant 3: Matching dictated name to written word across all three word sets (CA)

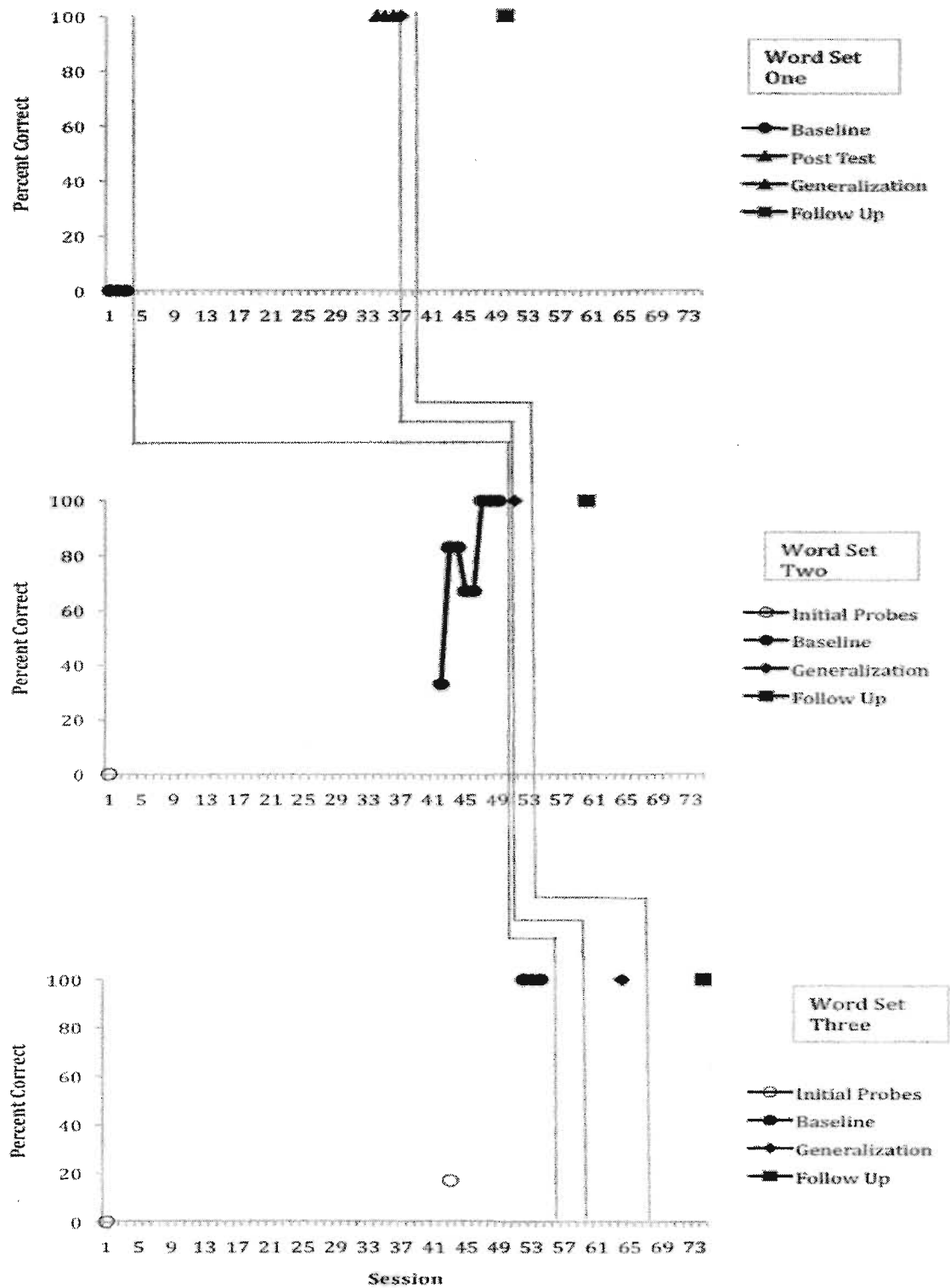


Figure 24: Participant 3: Oral naming of written word across all three word sets (CD)

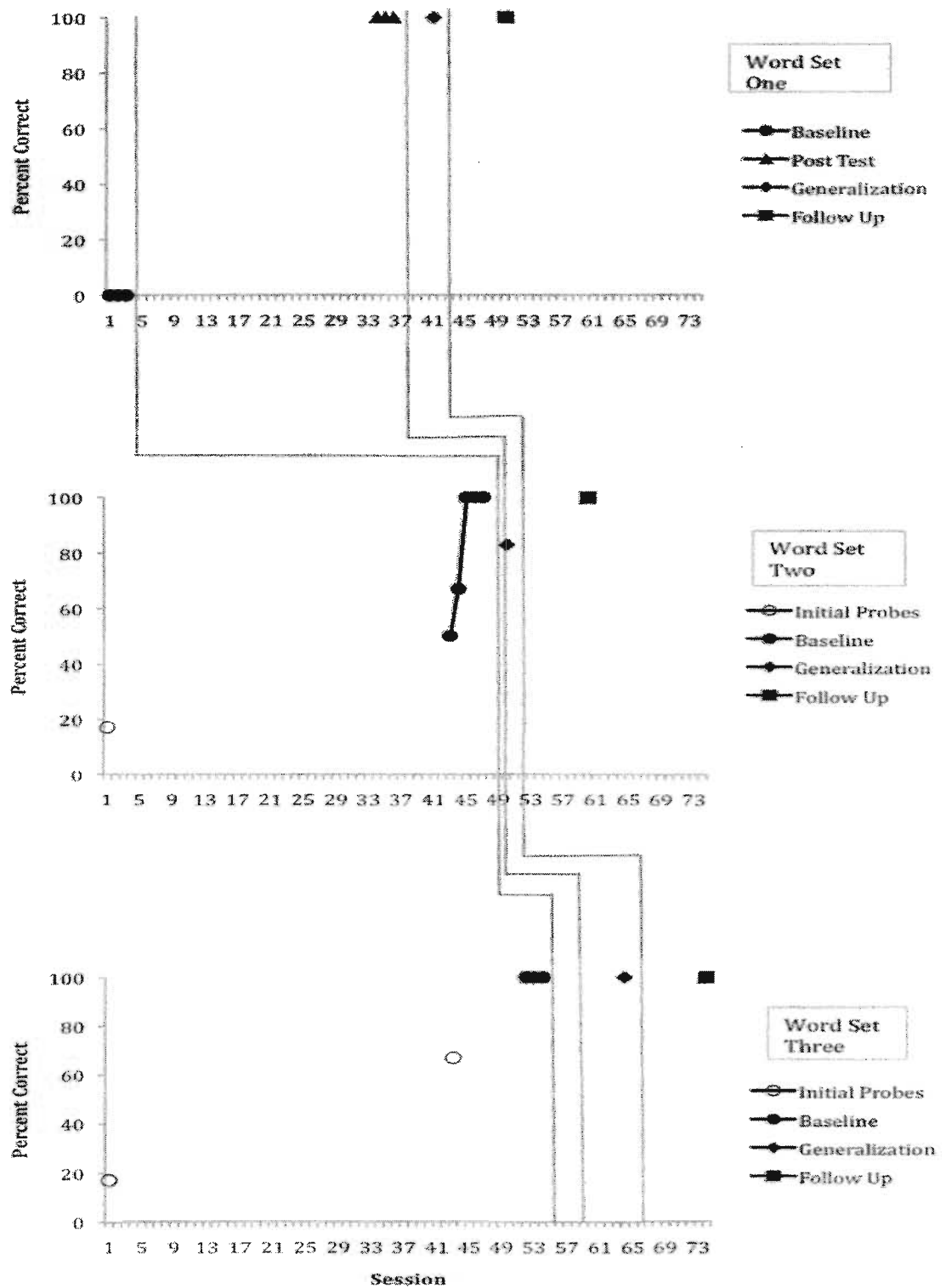


Figure 25: Participant 3: Matching picture to written word across all three word sets (BC)

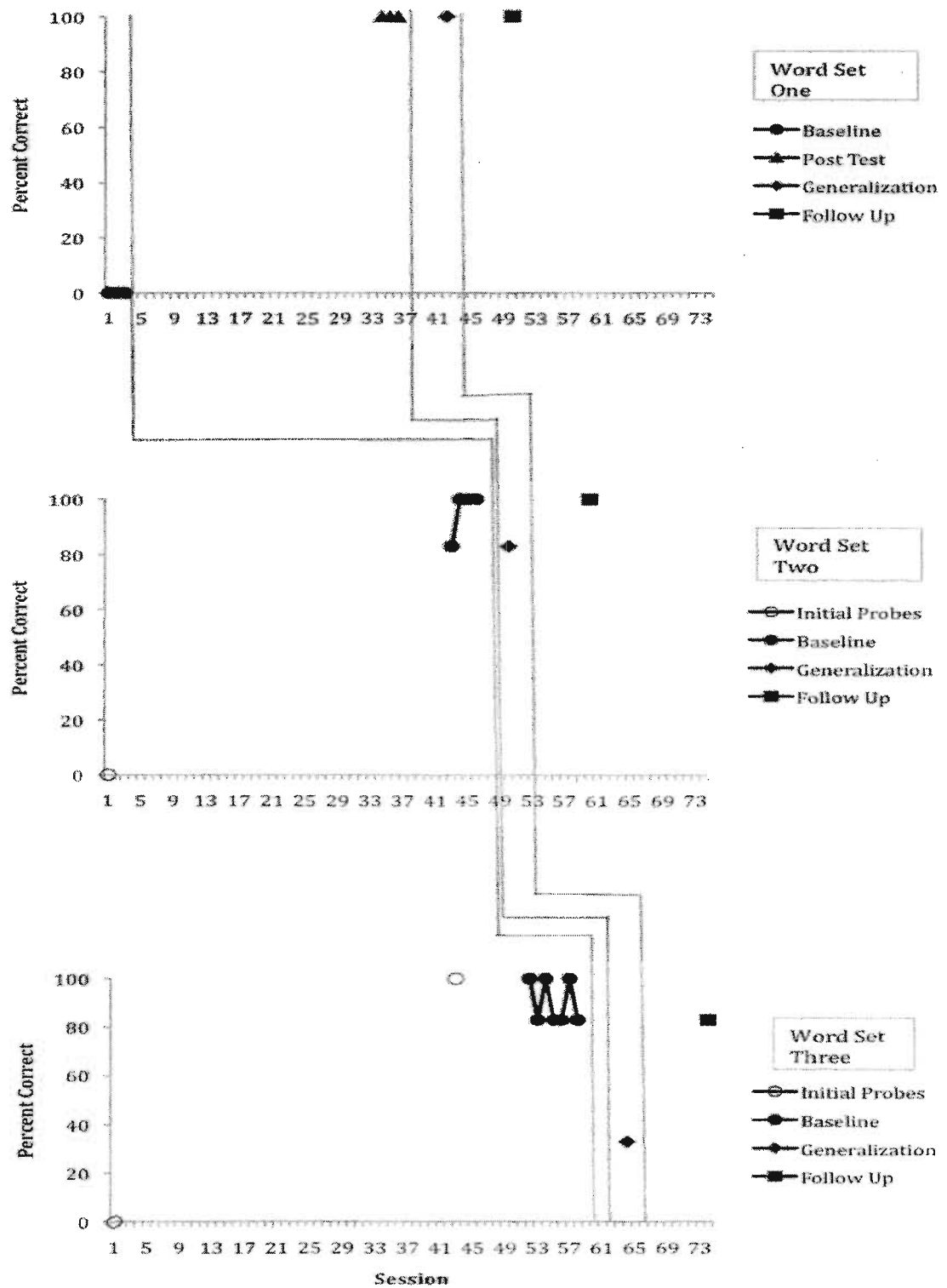


Figure 26: Participant 3: Matching written word to picture across all three word sets

Following baseline, Henry's Word Set One entered into the intervention phase. The upper graph of Figure 21 shows that a total of 23 intervention sessions, (a total of 690 trials) were completed for Henry to meet the mastery criterion. Gestural prompts were used throughout the intervention phase for Henry. For Henry, two words reached mastery at a faster rate than the third word. All three words were continued until the third word reached the mastery criterion.

Following the intervention phase, Henry's Word Set One entered the post test phase. For part 1 (combined test for symmetry and transitivity), Henry demonstrated 100% accuracy on the review tests (AB and AC) , which proceeded the combined tests. For the combined tests, Figures 26 and 25 show 100% accuracy on matching written words to pictures (CB) and matching pictures to written words (BC). Part 2 entailed testing in oral naming of written words (CD). Henry demonstrated perfect performances on the reviewed relations (BD and BC). Figure 24 shows that, following the review tests, Henry's performance for the oral naming of written words (CD) achieved criterion immediately with 100% across three sessions. Part 3 reviewed the matching of written words to dictated names (AC), which was maintained at 100%. This was then followed by testing the symmetry, matching dictated names to written words (CA). For CA (Figure 23), Henry's performance was variable. Testing continued in an attempt to stabilize the data and accuracy scores eventually dropped down to 0%. Part 3 testing was discontinued after 81 trials and Henry was then moved onto testing of Part 4. Part 4 began with a review of matching pictures to dictated names (AB) followed by the symmetry of that relation, matching dictated names to pictures (BA). Henry demonstrated perfect performance for the skill being reviewed (AB). For the symmetrical test however, Figure 22 shows that Henry's performance for the BA task was 44%, 67%, and 67% for the three data points. Note, following the completion of the third word set, post tests 3 and 4 were re-administered for Word Set One. At this time, significant improvements were

noticed, with performances of 100%, 56%, and 100% for the CA task and 100%, 100%, and 100% for the BA task. Post test part 5 looked at the reflexive relations AA, BB, and CC (Table 5), which were already discussed. Henry completed all post tests for word set one in a total of 600 trials across 19 days (including the re-test).

Following the post tests, a generalization session was completed. Figures 24, 25, 26, and 21 show that Henry's achieved 100% across all generalization probes except matching written words to dictated names (AC) where his performance was 83%. Henry's results indicate that his emergent performances were also generalized to novel stimuli.

When Henry's Word Set Three entered baseline, a follow-up probe was conducted for Word Set One. Figures 24, 25, 26, and 21 show that Henry's performance for oral naming of written words (CD), matching pictures to written words (BC), matching written words to pictures (CB), and matching written words to dictated names (AC) were maintained at follow up. Henry's performance for the BA and CA tasks (Figures 22 and 23) were both 33%, scores not far below the scores during post test.

Word Set Two. The words bus, fork, and bear were denoted as Word Set Two. Results of Henry's Word Set Two for each phase of the study (including the single probe session when Word Set One was in baseline) will be discussed next.

Figure 21 shows that Henry's performance for the skill that was taught, matching written words to dictated names (AC), was 0% at the initial probe. For the symmetrical relations, matching dictated names to pictures (BA) and matching dictated names to written words (CA) Henry's accuracy at the initial probe was 33% (Figures 22 and 23). Figures 24 and 26 show that the remaining relations, oral naming of written words (CD), and matching written words to pictures (CB) were at 0% at the initial probe, and Figure 25 shows that matching pictures to written words (BC), was 17%.

In order to obtain a stable baseline for Henry's second word set across all the relations, 254 test trials were completed. Interestingly, Figure 21 shows that Henry's performance for matching written words to dictated names (AC) was at 100% in baseline, prior to the intervention phase, and despite the skill not being in his repertoire during the initial probes. Henry's accuracy scores on the matching of dictated names to pictures (BA) ranged from 22% to 44% (Figure 22) and for matching dictated names to words (CA) scores ranged from 33% to 44% (Figure 23). The middle graphs in Figures 24, 25, and 26 show that oral naming of written words (CD), matching pictures to written words (BC), and matching written words to pictures (CB), were all initially low, but increased to 100% across the last three sessions.

Following baseline, Henry's Word Set Two was to enter into the intervention phase. However, given that Henry's baseline responding for matching written words to dictated names (AC) was already at mastery, as were many of the relations, the intervention and post test phase for Word Set Two were not completed. The exception to this was the repeat of post tests 3 and 4 for Word Set Two. Following the completion of the third word set, post tests 3 and 4 were administered for Word Set Two. At this time, significant improvements were noted, with performances of 89%, 100%, and 78% for the CA task and 78%, 100%, and 89% for the BA task.

Following the baseline, a generalization session was completed. The middle graphs of Figures 24, 25, 26, and 21 show that Henry achieved 100% for the generalization probes for oral naming of written words (CD) and matching written words to dictated names (AC). Performance for matching written words to pictures (CB) and for matching pictures to words (BC) was nearly perfect with performance at 83%, indicating his emergent performances were generalized to novel stimuli.

When Henry completed the generalization probes for Word Set Three, a follow-up probe was conducted for Word Set Two. Figures 24, 25, 26, and 21 show that oral

naming of written words (CD), matching pictures to written words (BC), matching written words to pictures (CB), and matching written words to dictated names (AC) were maintained at follow up. For the BA and CA tasks (Figures 22 and 23), performance was actually significantly higher than it was in the original post tests, which is consistent with the results of the re-test already discussed. Namely, accuracy for both CA and for BA was 89%.

Word Set Three. The words duck, cat, and book were denoted as Word Set Three. Results of Henry's Word Set Two for each phase of the study (including the single probe sessions when Word Set One and Two were in baseline) will be discussed next.

Figure 21 shows that Henry's accuracy level for matching written words to dictated names (AC) was 0% at the initial probe and was 50% for the second probe. For matching dictated names to pictures (BA), and matching dictated names to written words (CA) Henry's performance for the first probe sessions was at 33% for both relations (Figure 22 and 23). For the second probes, results indicate 100% performance for matching dictated names to pictures (BA) and 78% for matching dictated names to written words (CA). Figure 24 shows that oral naming of written words (CD) was at 0% for the initial probe. For the second probe, accuracy on oral naming of written words was 16.7%. Figure 25 shows that matching pictures to written words (BC) was 16.7% for the first probe session and for the second probe session was 67%. Lastly, Figure 26 shows that for matching written words to pictures (CB) Henry's performance was 0% for the first probe session and 100% for the second probe session.

In order to obtain a stable baseline for Henry's third word set across all the relations, 315 test trials were completed. Henry's performance on the skill to be taught, matching written words to dictated names (AC) was near perfect during baseline (Figure 21). Henry's performance on matching of dictated names to pictures (BA) was initially low, with a score of 33% but gradually increased to 89% (Figure 22). Similarly, accuracy

for matching dictate names to written words (CA) was variable between 33% to 83% (Figure 23). Baseline was stopped when the final three sessions were 83%, 67%, and 56%. Figures 24 and 25 show that Henry's performance for oral naming of written words (CD) and matching pictures to written words (BC) was perfect. Figure 26 shows that accuracy scores for matching words to pictures (CB) ranged from 83%-100%.

Following baseline, Henry's Word Set Three was to enter into the intervention phase. However, given that Henry's baseline responding for matching written words to dictated names (AC) was already at mastery, as were most of the relations, the intervention and post test phase for Word Set Three were not completed.

Following baseline, a generalization session was completed. Figures 24 and 25 show that Henry's achieved 100% across both the CD and BC tasks for the generalization probes. Figures 26 and 21 show that performance for the CB task was only 33% and for the AC task, performance on the generalization probe was 67%.

Ten days following the generalization probe for Henry's Word Set Three, a follow-up probe was conducted for Word Set Three. Figures 24, 25, 26, 21, 22 and 23 show that all emergent performances were maintained at follow up.

Participant 4

Reflexivity. Table 6 shows the initial baseline, post test and follow up results for the identity matching tasks (AA, BB, and CC). Pat's performance at the initial baseline probes, across all three word sets, for matching names to names (AA) was and for matching pictures to pictures (BB) was near perfect for each word set. For matching words to words (CC) Pat's performance at the initial baseline probes was 33% for the first word set, 0% for the second word set and 33% for the third word set. For post test and follow up phases, Pat's reflexivity test scores for Word Set One ranged from 93% to 100%. For Word Sets Two and Three reflexivity scores are not available for post test or follow up as Pat's participation was discontinued after the first word set.

Table 6:

Participant 4: Reflexivity scores for all three word sets

Word Set	<u>Name to Name (AA)</u>			<u>Picture to Picture (BB)</u>			<u>Word to Word (CC)</u>		
	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up	Baseline	Post Test	Follow Up
1	83%	93%	100%	100%	100%	100%	33%	93%	100%
2	100%	n/a	n/a	100%	n/a	n/a	0%	n/a	n/a
3	100%	n/a	n/a	100%	n/a	n/a	33%	n/a	n/a

Word Set One. The words lion, two, and book were denoted as Word Set One.

Results of Pat's word set one for each phase of the study will be discussed next, and the results of the single probe sessions for Word Sets Two and Three will be discussed later when presenting the results for these word sets.

In order to obtain a stable baseline for Pat across all the relations, 291 test trials were completed. Figure 27 shows that matching written words to dictated names (AC), the skill that was taught, was 0% during baseline. Similarly, Figures 28 and 29 show that matching dictated names to pictures (BA) and matching dictate names to written words (CA) were stable at 0% following some initial variability. Figure 30, 31, and 32 show that oral naming of written words (CD), matching pictures to written words (BC) and matching written words to pictures (CB) were all at 0%.

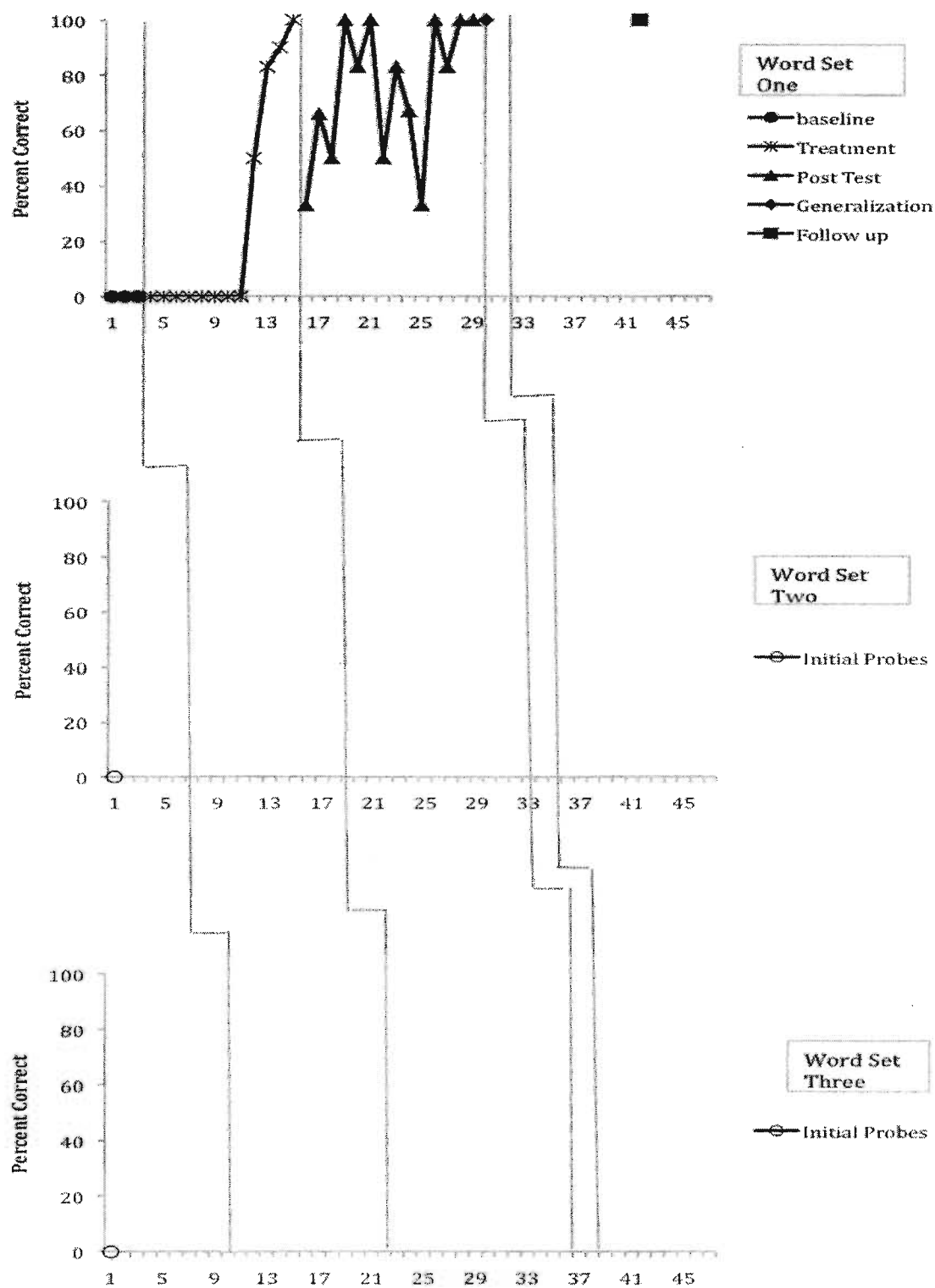


Figure 27: Participant 4: Matching written word to dictated name across all three word sets (AC-taught relation)

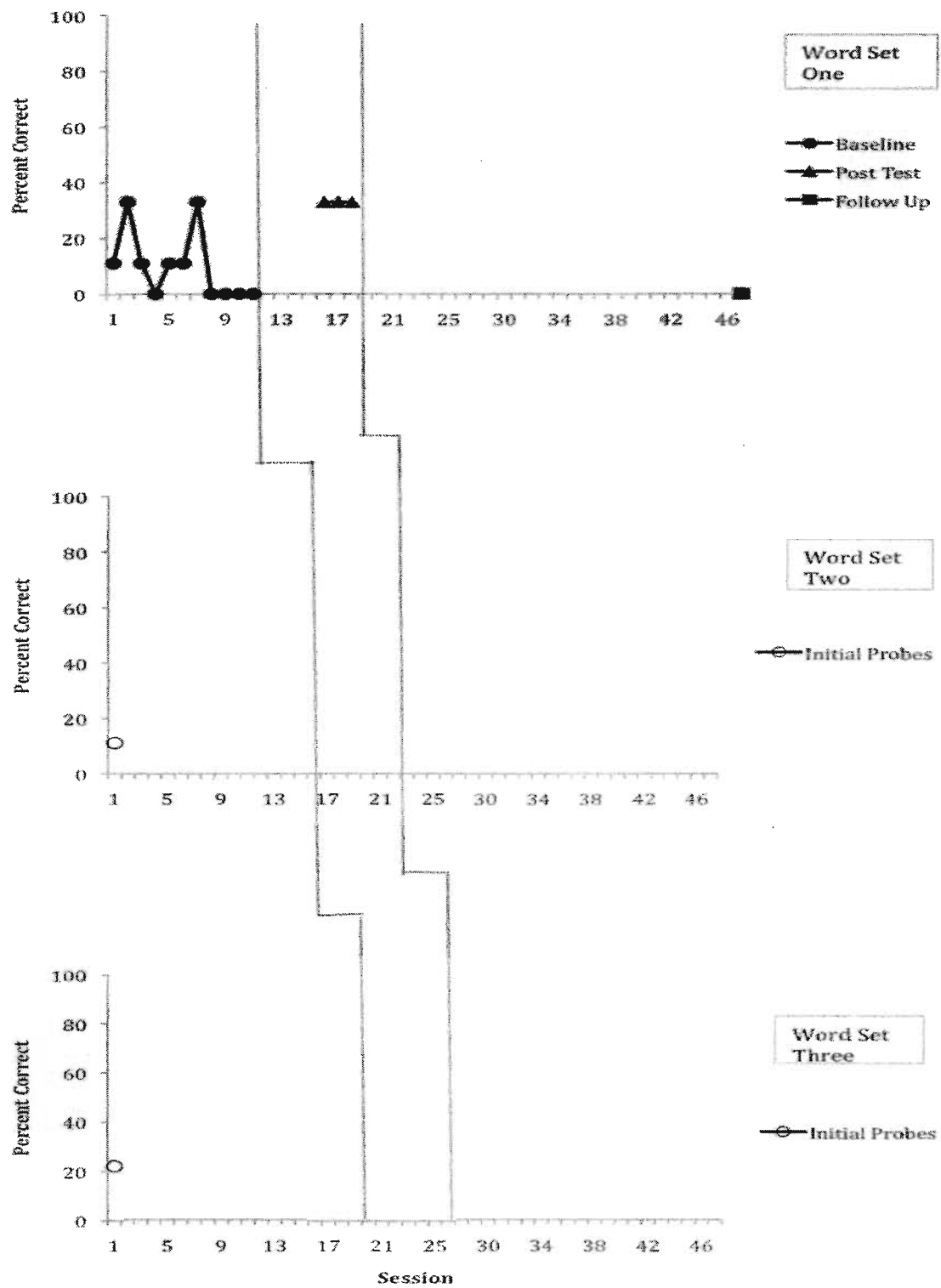


Figure 28: Participant 4: Matching dictated name to picture across all three word sets

(BA)

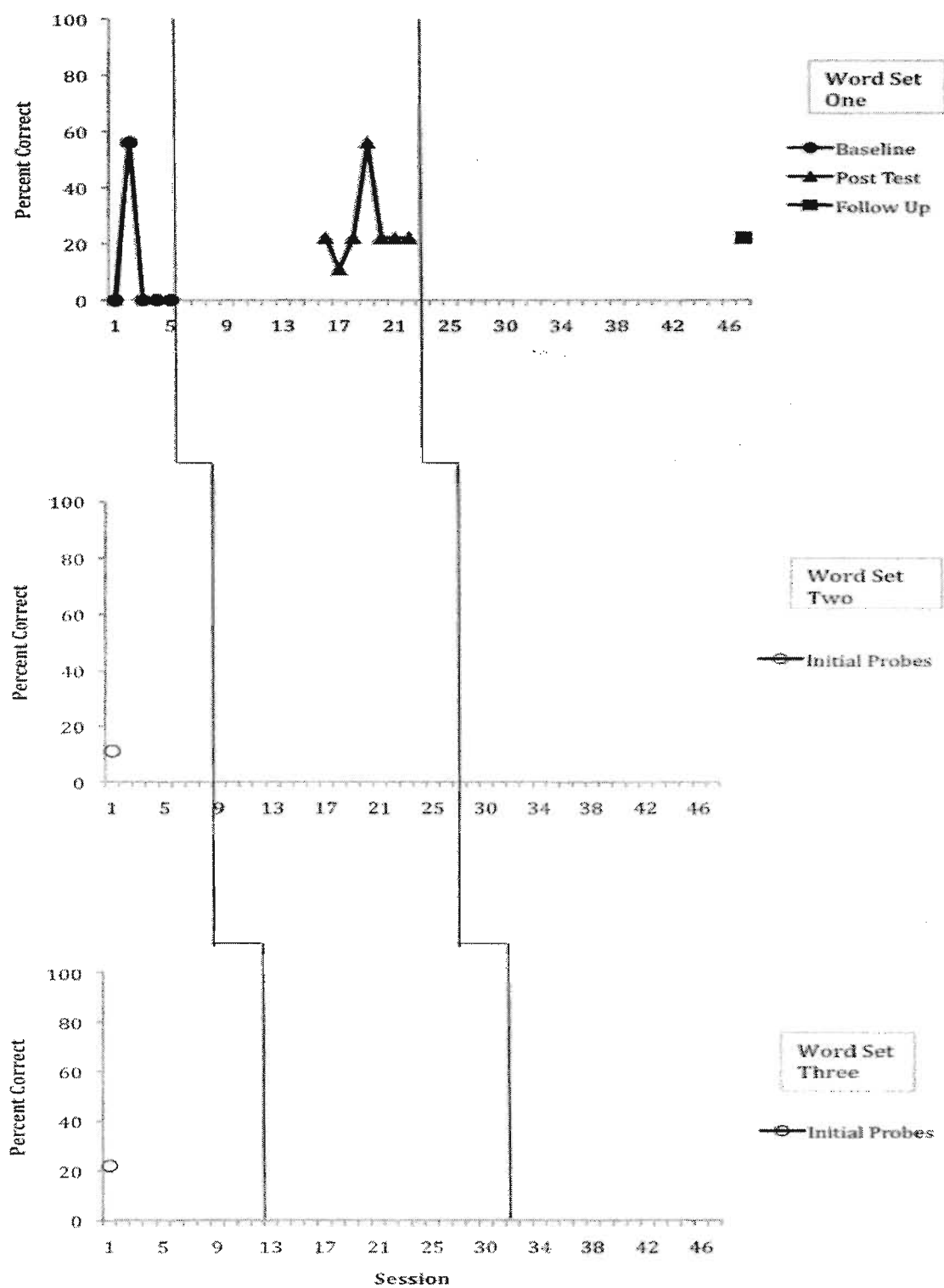


Figure 29: Participant 4: Matching Dictated name to written word across all three word sets (CA)

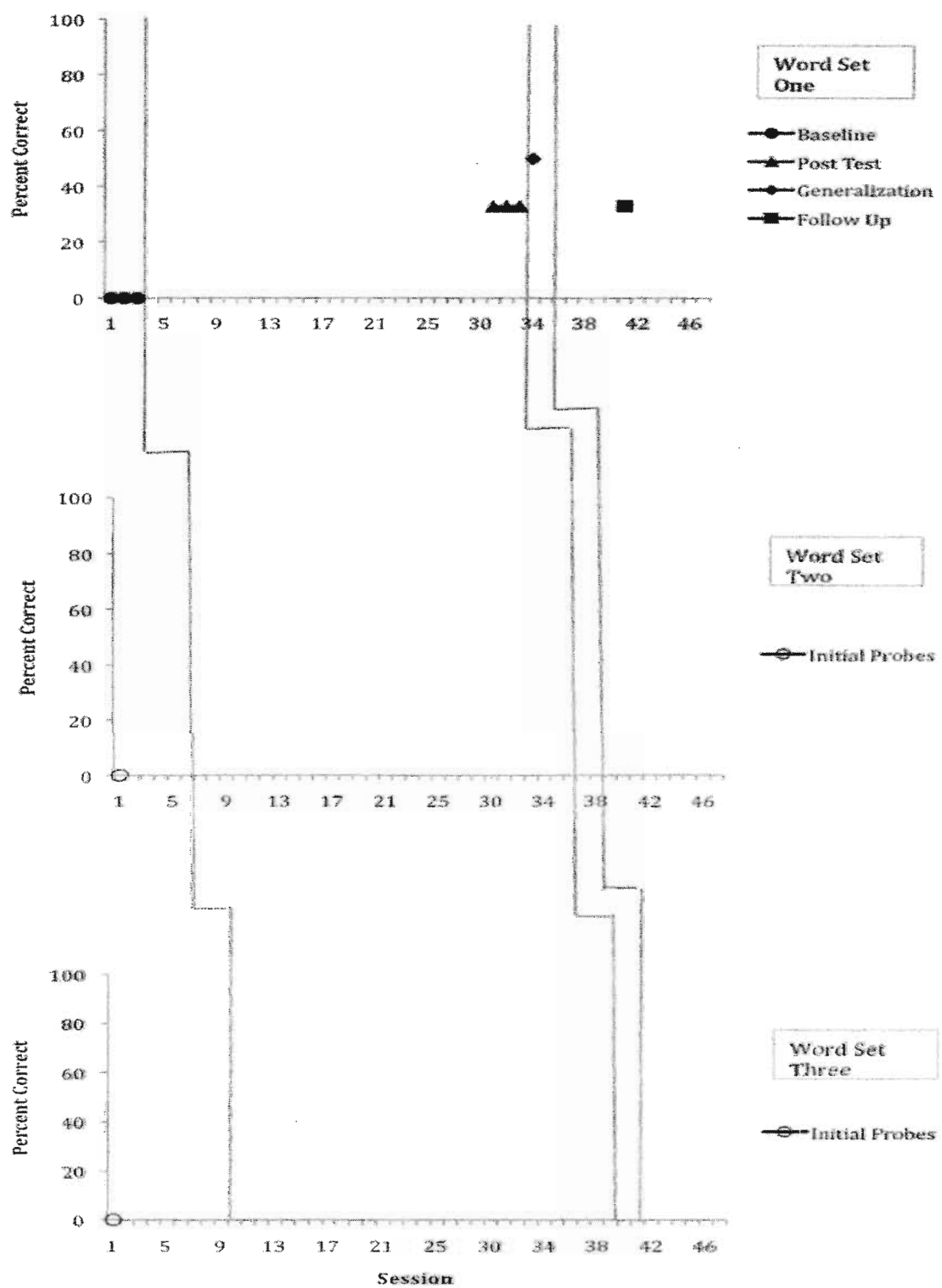


Figure 30: Participant 4: Oral naming of written word across all three word sets (CD)

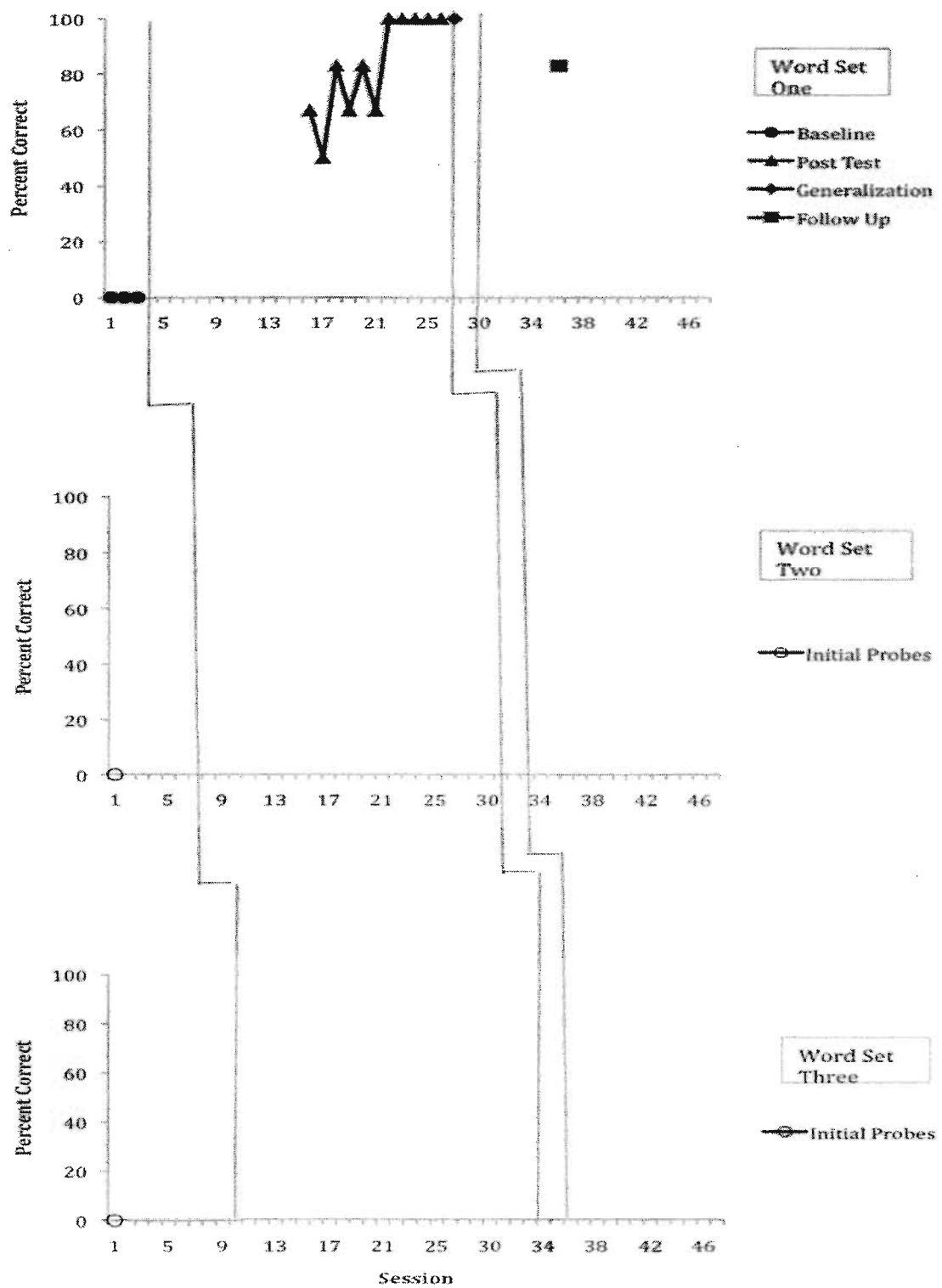


Figure 31: Participant 4: Matching picture to written word across all three word sets (BC)

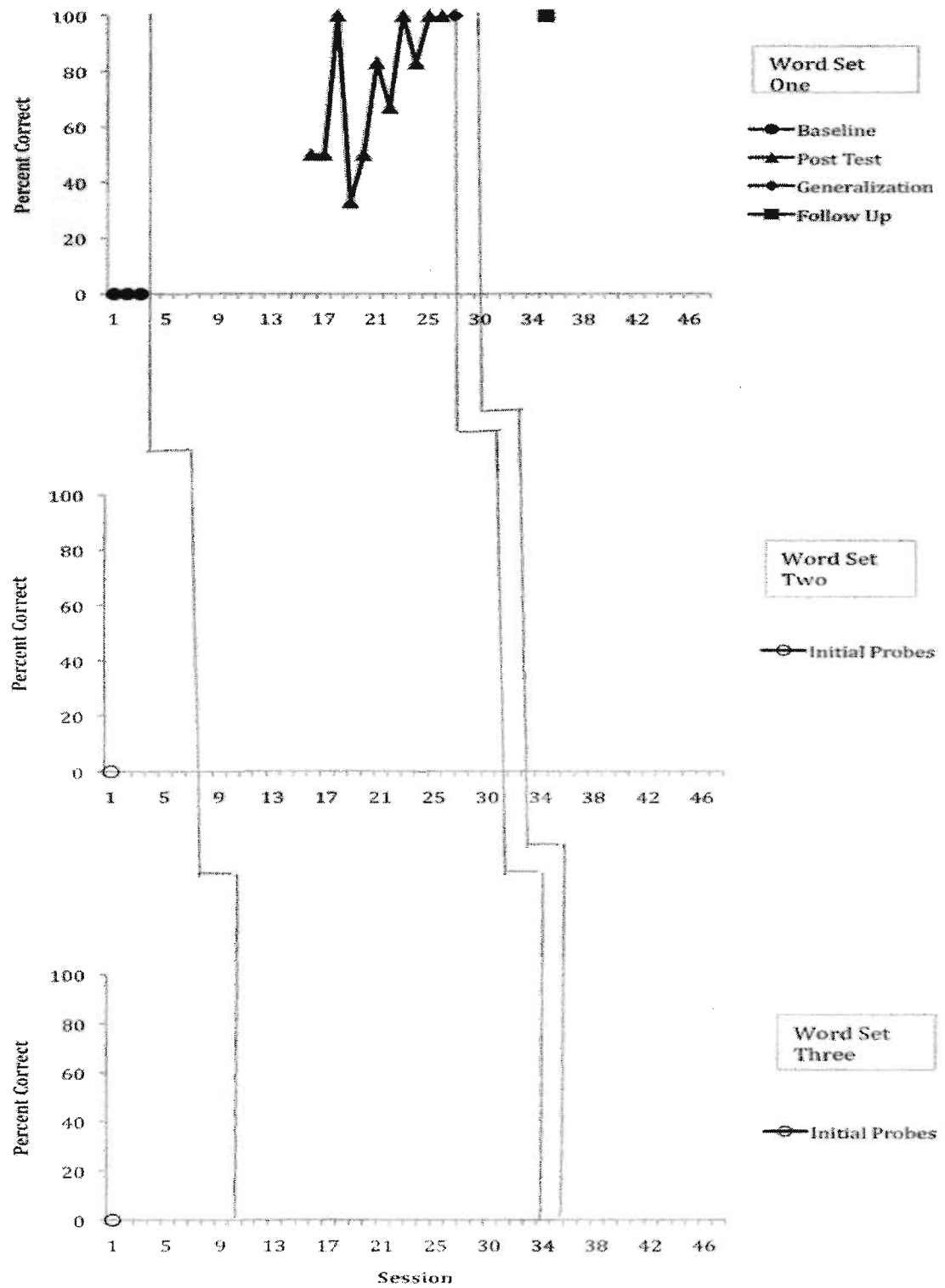


Figure 32: Participant 4: Matching written word to picture across all three word sets (CB)

Following baseline, Pat's Word Set One entered into the intervention phase. Figure 27 shows that a total of 12 intervention sessions, (a total of 360 trials) were completed for Pat to meet the mastery criterion. For Pat's first seven training sessions gestural prompts or positional prompts were used to prompt the correct response, however an increase in errors was reported as therapists attempted to fade their prompts. Prompting was therefore switched to a simultaneous prompting procedure (already described) using three types of prompts simultaneously; gestural prompts, positional prompts, and a verbal prompt referred to a spelling prompt. Mastery criterion for all three words was achieved at the same time.

As was already described Pat's original (one session only) post test was run in the order initially outlined in the methodology. The initial post test results were therefore gathered across one data point, and a total of 42 trials. Results from this initial post test are displayed on Figures 27-32 but are not discussed here as the procedure was modified for the next presentation and post tests were all restarted. The revised post test, as has already been discussed, contained five parts. It is important to mention that, as was already described, a correction procedure was included for Pat in which his behavior was corrected if his selection of the response was done in an inappropriate manner. In cases where this occurred the instructors would record a -/+, +/+, -/- or +/- indicating that a correction took place for Pat's behavior. In the events where this occurred the second data point was what was included for graphing purposes (i.e. if out of 6 trials 1 was recorded as -/+ the + was counted and the column was reported as 100%). For Part 1 (combined test for symmetry and transitivity), Pat's performance on the first review test (AB) was 83% for the first session and 100% across all subsequent sessions. Pat's the second review test (AC) was initially quite variable, but eventually (with the modified procedure) was relatively stable at 83% (i.e. 1 error) to 100% across three sessions (Figure 27). For the combined tests, which followed the review, Figure 32 shows that

matching written words to pictures (CB) was initially variable but (with the modified procedure) Pat achieved criterion on this test with scores of 100%, 83% and 100% across the three sessions. Similarly, Figure 31 shows that for matching pictures to written words (BC) performance was also quite variable but following procedural modifications was stable at 100%. Part 2 entailed testing in oral naming of written words (CD). Pat demonstrated 100% accuracy on the reviewed relations (BD and BC). Figure 30 shows that, following the review tests, Pat's performance for the oral naming of written words (CD) stable at 33% with consistently correct response for the word "book" and consistently incorrect responses for both "Lion" and "Two". For Part 3, Pat demonstrated 100% accuracy on the relation being reviewed (AC). For the symmetry (CA) task, however (Figure 29), Pat's performance was initial variable stabilizing at 22%. Observation indicated that Pat was consistently matching the sample stimulus to the comparison stimulus in the middle position. For Part 4, Pat's performance on the relation being reviewed (AB) was 100% across all sessions. For the BA task (Figure 28) however, Pat's performance was stable at 33% and observation again indicated consistent patterns of responding based on positional preferences. Finally, post test Part 5 examined Pat's performance on the reflexive relations AA, BB, and CC (Table 6), which was already discussed. With the revised order of the post test Pat completed all equivalence tests in a total of 579 trials across 15 days.

Following the post tests, a generalization session was completed. Figures 30, 31, 32, and 27 show that Pat achieved 100% on all generalization probes, except for oral naming where he achieved 50%.

At this point, a decision was made by the clinical team and the research team to discontinue Pat's participation in the study following the completion of Word Set One. The reason for this decision was based both on Pat's results from the study as well as the significant challenges in maintaining his motivation. While his first word set was

completed in 38 days (the same amount of time as Participant 1), significant challenges were had in achieving these results and the clinical team felt that his results overall did not warrant continuations. While Pat did eventually pass the equivalence tests for BC and CB, he failed to acquire CD, BA, and CA. This paired with the amount of time (and difficulty) taken to get to this point raised questions about the efficiency of this methodology for Pat. Pat's participation in the study was discontinued such that the clinical team could explore more efficient methods to teach him to read.

Pat's final phase in the study was therefore the follow up probe for Word Set One, which was completed 7 days following his generalization probes. Figures 31, 32, and 27 show Pat's performance for matching pictures to written words (BC), written words to pictures (CB), and matching written words to dictated names (AC) and demonstrates that the emergent performances were maintained at follow up. Pat's performance for the BA and CA tasks (Figures 28 and 29) were 0% and 22% respectively at follow up. Finally, Figure 30 shows that Pat's performance for oral naming of written words (CD) was 33% at the follow up probe.

Rate of Acquisition and Efficiency

Including word selection, baseline, intervention, post-test, generalization and follow-up for all three word sets, Participants 1, 2, and 3 required a total of 72, 67, and 80 days, respectively to complete all phases of the study. Participant 4 discontinued participation after 38 days. In looking at the rate of acquisition for the skill that was taught, matching written words to dictated names (AC), the following results can be seen. Participant 1 mastered the first word set in 11 sessions, the second word set in 7 sessions, and the third word set in 4 sessions. Participant 2 mastered the first word set in 17 sessions, did not require teaching for the second word set because it met mastery criteria in baseline, and the third word set was mastered in 5 sessions. Participant 3 mastered the first word set in 23 sessions and did not require teaching for both the second and third

word sets as both achieved mastery during baseline. Participant 4 mastered the first word set in 12 sessions but participation was discontinued.

Discussion

Children in this study were able to demonstrate up to eight emergent relations (for each of the nine words) after being taught only one relation. These results support previous demonstrations of emergent relations and provide preliminary evidence for the use of an SE paradigm to teach reading and reading comprehension to children with Autism in IBI settings. In the following discussion, I will first review the answers to our research questions. Next, I will compare our findings to evidence already published in the literature, followed by a discussion of how the findings in the present study extend the literature to date. Last, I will review the strengths and limitations of this study along with recommendation for future research.

Can children with Autism demonstrate reflexivity?

Reflexivity was tested by evaluating the children's matching of names to names (AA), pictures to pictures (BB), and written words to written words (CC). All of the participants in the study demonstrated reflexivity after being taught to match written word to dictated name (AC). It is important to point out that many of the participants performed accurately on tests for reflexivity during the baseline phase. This is not surprising, as all of the participants were attending an IBI program and had previous learning experience with matching tasks. Given the presence of these skills at the start of the study for many of the participants, one cannot conclude that these relations were the result of teaching the trained relation. This is consistent with what has previously been reported in the literature. Green and Saunders (1998) suggest that reflexivity test results are likely the result of pre-existing identity matching abilities, rather than the result of the experimental contingencies. They suggest testing for reflexivity only after the other arbitrary relations have been established rather than testing during baseline as well. For

two of the four participants however, matching written word to written word (CC) was not in their repertoire at the start of the study, but rather emerged following the acquisition of matching written word to dictated name (AC). For these two participants one can conclude that this relation was the result of teaching the trained relation.

Can children with Autism demonstrate symmetry and transitivity of printed words, pictures and dictated words?

In this study, symmetry and transitivity were evaluated by testing for the emergence of matching pictures to written words (BC) and matching written words to pictures (CB). Symmetry was further evaluated by testing for the emergence of matching dictated names to pictures (BA) and matching dictated names to written words (CA), using a language master card reader to provide the audio cues. All four participants in the study passed the combined tests for symmetry and transitivity, as evidenced by the emergence of matching picture to written word and matching written word to picture. For the additional symmetry tests however, only two of the four participants were successful. More specifically, only two participants correctly matched the audio cues from the language master to the corresponding pictures or written words. Recall that Sidman (1994) argues that when both auditory and visual stimuli are involved in the paradigm direct tests for symmetry are not possible, at least without changing the standard procedure involving simultaneous presentation of comparison stimuli. In fact, there is no evidence of published literature in which these tests of symmetry have even been attempted in cases where auditory and visual stimuli are included.

The indirect attempts to test for this symmetry were a modification of attempts made by Savona (unpublished, 2008). The fact that two of the four participants demonstrated these relations is important. The failure by the remaining two participants to demonstrate these relations appeared to highlight the limitations of the testing

paradigm rather than a lack skill. More specifically, the testing paradigm required the participant to remember the location of the card corresponding to the target audio cue.

In this study, the evaluation of the symmetry of the audio-visual relations involved the successive presentation of the auditory stimuli. It is important to note that this differs from typical methodology in stimulus equivalence in which comparison stimuli are all presented simultaneously (Sidman, 1994). Future research should examine whether successively presented stimuli can still provide a measure of symmetry, or whether modifications can be made to the testing paradigm used in this study, such as providing pre-training in the use of the language master.

Does oral naming of written words emerge following the teaching and emergence of the relations described above?

While from a theoretical perspective oral naming is not a necessary outcome to conclude that equivalence has occurred, from a clinical perspective the demonstration of oral naming is of great significance. Oral naming of written words is often taken as proof of reading from an educational perspective, even in the absence of the underlying comprehension. If by teaching a child to match written words to their dictated names (AC) if the child can not only match words to pictures and pictures to words (BC and CB) but can also read the words aloud, then the outcome has great clinical significance. Oral naming of written words emerged for the three children who completed the study, providing further rational for the potential clinical utility of this procedure.

Do emergent relations and the relation taught generalize to novel stimuli not presented during the baseline or post test phases?

Of the four participants in this study, all demonstrated overall good generalization to the novel stimuli presented in the generalization phase, with some variability. All four demonstrated generalization for the skill that was taught, matching written word to dictated name (AC), although Participant 2 did show some drop in

performance for this skill for Word Set Two. All four participants demonstrated generalization for matching picture to written word (BC) for all three word sets. Interestingly however, two of the participants showed less generalization (but still within an acceptable range-72%) for the symmetrical relation, matching written word to picture (CB). This is surprising given that the novel stimuli used for BC and CB were identical, and both of these participants demonstrated BC and CB equally within the post test phase. For oral naming (CD), all four participants demonstrated generalization scores that were on par with their scores during the post test phase (i.e. Participant 4 did not pass the tests for oral naming during post test and did not pass the tests for oral naming during generalization either).

Are emergent relations maintained over short durations of time (i.e. 7-10 days)?

Mixed results for the follow up phase were seen across the four participants. All four participants demonstrated maintenance of the relations at the follow up probes for the reflexive relations (AA, BB, CC) and the relation that was taught (AC). Similarly, all 4 participants demonstrated that matching pictures to written words and matching written words to pictures (BC and CB) were maintained at follow up. Of the three participants that demonstrated oral naming of written word (CD) during the post test (Participants 1-3), one of the participants (Participant 1) did not demonstrate maintenance of this skill. That participant transitioned to school at the end of study and follow-up probes occurred in his home setting. This could have impacted performance, as he experienced a greater delay between the generalization and follow up phases, and the follow probes were run with his siblings in the room often causing distraction. Both Participants 2 and 3 however did successfully maintain the oral naming during follow up phase. For the last two relations, BA and CA, mixed results were observed. Generally speaking, the two participants who did not demonstrate these skills during post test (Participants 2 and 4) also did not demonstrate them at follow up. The two participants who did demonstrate

these skills during post test maintained these skills to follow up, however, there was some drop in correct responding from the post test phase. Both participants maintained scores at 70% or higher but neither one achieved 100% maintenance of these relations.

Note that the study did not include a systematic maintenance phase as would be typical within an IBI program. Namely, when participants finished the generalization phase they did not see the stimuli again for seven to ten days at which point the follow up probe was completed. This is in sharp contrast to what would be typical within their IBI programs where a systematic maintenance schedule would be implemented to enhance generalization over time (e.g. children would go from seeing a program every day, then every other day and so on). The lack of a systematic maintenance program may explain some of the mixed results at follow-up. However, it is important to note that despite some minimal drop off for BA and CA, and the lack of maintenance for CD for Participant 1, the majority of emergent relations were maintained at follow up.

Do children with Autism acquire the taught relation (i.e. matching written word to dictated name) more quickly, following the demonstration of emergent relations on a previous word set?

This final research question is not a simple one to answer, given some of the unique outcomes that resulted during this study. Namely, the acquisition in the absence of teaching that occurred resulted in the intervention phase being unnecessary for one word set for Participant 2 and for two word sets for Participant 3 (because the skill to be taught achieved mastery during baseline). This makes an evaluation of their rates of acquisition more challenging. In looking at the results however, the following conclusions can be made. Overall it would appear that rate of acquisition increased on subsequent word sets. For Participant 1 this is very clear. For Participant 2 days to acquisition decreases from 17 sessions for the first word set to five sessions for the third word set (word set two was at mastery during baseline), which is a clear increase in his rate of acquisition. Participant

3 took 23 sessions and did not require an intervention phase for second and third word set due to mastery. On some level one could say that his rate of acquisition went from 23 sessions to zero sessions. Participant 4's rate of acquisition cannot be evaluated due to the termination of his participation in the study. The mechanism underlying this increase in speed of learning was not studied directly, but we hypothesize that this teaching paradigm increased student's attending to text, which may in turn increase motivation for these tasks. It is possible that the child increased their attending to stimuli in their environment after exposure to the teaching situation.

Comparison to Other Studies

Replication of Sidman's original study. In the present study, four participants who could already match pictures to dictated names and could name pictures were taught to match nine words (or 3 words for Participant 4) to dictated names. Following the training and without any additional teaching the participants were all able to match written words to pictures and match pictures to the written words. Three of the four participants were also able to name the written words. This is consistent with Sidman's original study, where a subject who could already match pictures to their dictated names and could name pictures was taught to match 20 written words to dictated names and without any additional teaching, was able to match written words to pictures, match pictures to written words and name the written words (Sidman, 1971). The favorable outcomes on tests for equivalence found in the present study are comparable to other results found in the literature. Melchiori and colleagues (2000), for example, also found participants were able to match written words to pictures, match pictures to written words, and to read the words, after being taught to match written words to dictated names and to construct the printed words.

Gradual emergence. In the present study, two of the four participants did not demonstrate one or more of the equivalence relations right away but rather did so only

after a number of repeated trials. This is consistent with what has been found in the literature. In fact, many authors have discussed gradual emergence of relations after repeated testing (e.g., Devany, Hayes, & Nelson, 1986; Driscoll & Kemp, 1996; Stromer et al., 1992). However, two of the participants demonstrated equivalence right away and for one of the participants where gradual emergence did occur, it did not occur for every relation. Given this information, and the fact that gradual emergence does not always occur, future studies may want to explore why some individuals demonstrate equivalence immediately, while others do so only after repeated testing.

Test order. The revised test order that was previously described is in line with the current literature. The test order moved the reflexivity tests to the end of the test sequence. Discussions regarding reflexivity in the literature have hypothesized possible interference with equivalence. More specifically, Green and Saunders (1998) suggested that reflexivity tests may cue the participants to identify commonalities between stimuli presented in MTS trials as a result of having completed identity matching trials where the search for the similar features is the goal. Saunders and Green (1992) suggested that reflexivity tests be conducted in separate sessions from the other relations being tested and later suggested that reflexivity tests be conducted last (Green & Saunders, 1998). While several different options regarding retesting and test order have been and are being explored in the literature, in the present study the revised test order was similar to that of Sigardottir and colleagues (1990) and involved strategic testing by first reviewing relevant trained (or known) relations and then testing for the emergent relations (for example, review AB and AC then test BC and CB).

Absence of oral naming in Participant 4. In the present study, three of the four participants demonstrated oral naming of the written words but Participant 4 did not. Mixed results with respect to oral naming is not uncommon in the stimulus equivalence literature. In fact, in a recent article Sidman (2009) noted that after learning to match

written words to their dictated names, some subjects named the written words while others were not. Given the mixed results in the literature with respect to oral naming future studies may want to explore why this relation emerges in some individuals and not others.

Extending Other Studies

First applied study of stimulus equivalence to teach reading to children with Autism. While there have been a few applied studies of stimulus equivalence (Driscoll & Kemp, 1996; Eikeseth & Jahr, 2001; LeBlanc et.al., 2003; Mackay, 1985), the only applied study to date to teach reading to children with Autism is a study completed by Eikeseth and Jahr (2001). This study however only tested for symmetry and therefore cannot be said to have been a complete study on stimulus equivalence. The present study in contrast included tests for reflexivity, symmetry, and transitivity thereby meeting Sidman and Tailby's (1982) criteria for demonstration of an equivalence class. In doing so, the present study is the first applied study of using stimulus equivalence to teaching reading skills to children with Autism. This study therefore provides a demonstration of stimulus equivalence in the real world. In addition to being an applied study, this study also demonstrates the field effectiveness of stimulus equivalence in that the interventions were implemented in real IBI classrooms, with the children's regular clinical teams, under less than ideal situations. This is in contrast to much of the applied research to date, where the interventions are implemented by graduate students, often in more controlled settings. The successful demonstration of equivalence for three of the four participants under these conditions provides support for the applicability of stimulus equivalence to real world teaching.

Acquisition in the absence of teaching. A unique result, not found in the previous literature, occurred in the present study. More specifically, in three of the four participants an outcome occurred which for the present purposes will be called

acquisition in the absence of teaching. More specifically, the study demonstrated that the relations tested were not in the children's repertoires during the initial baseline probes. However, in a number of cases, after learning one or two word sets, emergent relations occurred without the child being taught the target skill (matching written word to dictated name). For Participant 1 for example, performances were low for the relations involving the third word set at the initial probes and the second probes. After completing the first and second word sets, performance during the baseline for the third word set demonstrated many of the relations were now already in the child's repertoire. Following a brief teaching period of AC (to increase performance to 100%) all of the relations emerged. Similar results were obtained for Participant 2 for the second word set as AC (and many of the other relations) were at mastery during the baseline phase, despite not being in his repertoire at the probes. For Participant 3, this effect was even more significant in that he acquired both the second and third word sets after completing the first word set. While it can not be known for sure at this time the cause of this effect, one possible explanation is that after learning one or two word sets that words themselves came to be meaningful as opposed to merely being arbitrary symbols. That, combined with the fact that the words being taught were common functional words, it is possible that the participants began to attend more to the stimuli in their environments, which may have played a role in the acquisition in the absence of teaching. Alternatively, it is possible that this result may have been due to previous exposure to the words or familiarity with them, despite attempts to control for this throughout the study.

This pattern of acquisition in the absence of teaching has not been found in previous literature. The majority of previous research has used a pre-post design, unlike this study. The multiple probe design used in this permitted the detection of these acquisition effects. A pre-post design would likely have tested all nine words at the same time and therefore would not show if acquisition in the absence of teaching had occurred.

Future studies should be conducted replicating the multiple probe design in order to determine whether this effect was related to a variable specific to the present study or if in fact the benefits of teaching using a stimulus equivalence approach are actually greater than what has previously been considered. More specifically, if by teaching 1 or 2 skills is the result not only the emergence of many other skills with those stimuli but also this emergence with other stimuli as well?

Testing the symmetry of auditory-visual stimuli. In the present study, an attempt was made to test the symmetry of the relations involving both auditory and visual stimuli. This differs from other studies (e.g. Sidman, 1971) where the symmetry of the auditory-visual relations is inferred by the emergence of other relations. Two of the four participants in this study demonstrated the symmetry of the auditory-visual relations, providing a more complete demonstration of equivalence than would be possible if these relations were only inferred.

Strengths

Multiple probe design across word sets. In the present study, a multiple probe design (Horner & Baer, 1978) across word sets was used, which was similar to the research design used by Driscoll and Kemp (1996). The present study differs from Driscoll and Kemp (1996) in a couple of key ways. First, Driscoll and Kemp (1996) taught participants two word sets, whereas the present study attempted to demonstrate control across three word sets. In addition, Driscoll and Kemp (1996)'s study included a multi-treatment method, comparing two different teaching sequences. In the present study, the multiple probe design was implemented for a consistent teaching and testing sequence across all participants. Furthermore, the use of a multiple probe design differs from most (if not all) other studies of stimulus equivalence. The majority of studies utilize a pre-post design where relations are tested prior to the teaching phase and after the teaching phase (e.g. De Rose et al., 1996; Leblanc et al, 2003). The advantage of the

multiple probe design utilized in the present study is that it allows some demonstration of control that the changes seen from pre-test to post-test are in fact the result of the stimulus equivalence paradigm and not another extraneous variable.

Absence of reinforcement. In the present study, reinforcement was not delivered for the target responses during any of the testing phases. Instead, an interspersal procedure was used where mastered trials (unrelated to reading) were interspersed with testing trials and were reinforced, or the child was reinforced for appropriate attending skills (sitting, looking, etc). This differs from some other studies (e.g., Mackay, 1985; Melchiori et al., 2000) where reinforcement was delivered throughout testing and teaching. Absence of reinforcement or the use of interspersal procedures has been documented in the literature as well (Driscoll & Kemp, 1996; Sidman, 1994). Despite some of the mixed approaches in the literature regarding the use of reinforcement during the testing phases, the absence of reinforcement in the present study is an important variable. In refraining from reinforcing the test trials one can eliminate the explanation that the relations emerged due to teaching (i.e. reinforcement of correct responses) versus truly being emergent relations.

Six-choice match-to-sample procedure. According to Green and Saunders (1998), much of the literature on stimulus equivalence involves match-to-sample procedures in which two comparison stimuli are available on each trial. Green and Saunders (1998) argue however that procedures involving two-choice MTS procedures have many confounds and they recommend instead using three-choice MTS procedures to diminish these confounds. In the present study, a six-choice MTS procedure was utilized. One could hypothesize that with a six-choice MTS procedure the confounds present in the two-choice procedure are even further removed, providing additional confidence in the results that were obtained.

Generalization and maintenance. Two components that were included in the present study that are not frequently reported in the stimulus equivalence literature are the inclusion of both generalization and follow up phases. Driscoll and Kemp (1996), did include a generalization phase where they measured generalization from picture labeling to object labeling, but they did not test generalization for any additional relations. They also measured maintenance in their study, but again only measured maintenance for the trained skills. In the present study, four exemplars were included throughout teaching and testing phases and a novel fifth exemplar was introduced to test for generalization. In doing so, the study was able to demonstrate that after teaching children one skill, they not only demonstrated emergent relations with one exemplar, but also demonstrated generalization across a variety of exemplars. While often not included in studies, Sidman's initial study also used multiple exemplars in teaching in order to limit the possibility that the participant would not attend to some irrelevant aspect of the stimuli (Sidman 2009). These results are promising, given the information in the literature regarding children with Autism and difficulties with generalization (Lovaas, Koegel, & Schreibmann, 1979). Similarly, from a clinical perspective, for a skill to be considered mastered it must not only be generalized, but be maintained over time. Thus, the utility of stimulus equivalence as a teaching approach from a clinical perspective depends not only on the emergence of relations but also on the maintenance of these skills over time. The present study evaluated the maintenance of relations tested including reflexivity (AA, BB, and CC), symmetry and transitivity (BA, CA, BC, and CB) the taught relation (AC) and oral naming (CD). As was already discussed the results overall were favorable indicating that emergent relations were in fact maintained over time. The inclusion of both the follow up and generalization phases in this study further strengthen the results obtained.

Functional words selected for teaching. Another strength of this study, from a field effectiveness perspective is the functionality of the words being taught. Words selected differ from some of the previous literature (Devany et al., 1986; Eikeseth & Smith, 1992; Green, 1990), in that actual words, rather than pseudo words were selected for each child. In addition, because the clinical team were directly involved in the word selection for each child, the words taught were individualized and therefore relevant to the child.

Treatment Fidelity. An additional strength of this study is the inclusion of measures to attempt to control for issues of treatment fidelity. More specifically, IOA data collection and treatment integrity measures were included for 20-25% of all phases of the study. Very strong IOA and integrity results were reported with IOA above 90% for all phases (and in fact above 95% for most phases) and integrity at 97% or above for all phases. The high treatment integrity results are of particular importance, given that this study examined field-effectiveness. To be able to achieve 97% or higher integrity scores for procedures being delivered by front line staff (as opposed to masters level students who often implement procedures in research), in a busy IBI setting, is quite impressive. These results provide further support to the field effectiveness of stimulus equivalence.

Limitations

Absence of reinforcement. While the absence of using reinforcement in the testing trials is a strength of this study for the reasons outlined above, it can also be seen as a limitation. The failure of Participant 4 to demonstrate a number of the relations, and his eventually removal from the study were directly related to the absence of reinforcement during the testing phases. Sidman explains the dilemma of whether or not to use reinforcement best in saying “if we reinforce, seemingly successful transfer might be attributable to reinforcement rather than the subject’s training history. If we do not

reinforce, a seeming failure of transfer might be attributable to extinction during the test rather than an inadequate training history (Sidman, 1994, pg 188)". It would appear that the benefits of not reinforcing outweigh the limitations, however further exploration is needed in terms of how to modify procedures for participants who appear more dependent on response specific feedback.

Emergent relations following the revised test order. As was already discussed, the test order was revised when Participants 1 and 4 failed to demonstrate emergent relations. This presents a limitation to the results as Sidman would argue that if equivalence had been established the test results would be positive irrespective of the testing sequence (Saunders & Green, 1998). It has been documented in the literature however, that retesting under revised test orders following failed tests on equivalence have demonstrated positive results (Saunders, Wachter, & Spradlin, 1988; Sidman, Kirk, & Willson-Morris, 1985; Sigurdardottir et al., 1990; Vause, Martin, Marion, & Sakko, 2005). In addition, it has been discussed in the literature that the results of equivalence tests may be related to the order in which the relations are presented. Green and Saunders (1998), for example, suggested the possibility that a review of training trials prior to related testing trials might effect the results of the test trials. Similarly, Saunders and Green (1992) supported the hypothesis of the possible role that test trial order may have on the result of equivalence tests, suggesting that different test trial orders may result in different outcomes on tests for equivalence.

Acquisition in the absence of teaching . Similarly, while acquisition in the absence of teaching is a strength from a clinical perspective, in that the participants in this study acquired a number of skills with a number of stimuli with very little teaching, from a control perspective this effect can be thought of as a limitation. More specifically, the multiple probe design was utilized in order to demonstrate experimental control across word sets and increase confidence in the conclusion that the stimulus equivalence

paradigm was responsible for the results and not other extraneous variables. The acquisition in the absence of teaching that occurred in some ways minimizes this control. What can be said is that at the start of the baseline (i.e. for the initial probes) the skills were not in the children's repertoire, in addition, in many cases the skills were still not in the children's repertoire at the second probe. Therefore, some element of control has still been maintained but it is possible that other extraneous variables may be responsible for the emergence of the second and/or third word sets in Participants 1-3.

Limited control in field effectiveness research. Field effectiveness research lacks some of the controls possible in laboratory studies. The IBI environment does not control for the possibility of all confounding variables. In this study, for example, the teachers and family members involved with the participants were instructed to avoid teaching or exposing the child to the target words during the course of the study. There was no guarantee however, that the child did not encounter the words during the study, which provides some limitations to the results.

Small sample size. The sample size for this study was relatively small (4 participants) and due to the number of testing trials required, only nine words (3 sets of 3) were taught to each child. While this may limit conclusions, our sample size is within the guidelines of the criteria discussed by Chambless and Hollon (1998) with respect to conclusions of effectiveness and the results of this study provide some preliminary evidence into the field effectiveness of stimulus equivalence in teaching reading to children with Autism. Replications are needed however to provide further support for these results.

Future research

Efficiency and economy of teaching. Stimulus equivalence methodologies are hypothesized to offer an efficient teaching methodology where practitioners get a "lot of "bang for the buck", because after just a few conditional discriminations are established

through direct training, many others typically emerge “for free” (Green, 2001 p. 80). In fact, according to Sidman (2009), equivalence relations can foster productivity, where a small amount of work can lead to the emergence of new skills that were not taught. In the present study, three of the four participants passed the majority of the tests for equivalence. In doing so, stimulus equivalence has been demonstrated to be an efficient and economical teaching methodology where participants were taught to relate nine words to their corresponding written words and without any additional teaching were able to do up to 72 additional tasks involving the relations between pictures, text and spoken words. While these results are promising, it is important to remember that conclusions regarding the efficiency and economy of this approach are still preliminary. Efficiency is a relative term, which requires a comparison to another approach. We can say that it appears that stimulus equivalence is an efficient approach, but what is not known at this time is whether or not stimulus equivalence is more efficient than another approach. Future studies should compare rate of acquisition using a stimulus equivalence methodology with rate of acquisition in teaching each of the relations individually.

Replication to explore acquisition in the absence of teaching effects. As was already discussed, an interesting side effect occurred for three participants in this study where skills were acquired on subsequent word sets prior to teaching. Future studies replicating the multiple probe design of the present study, should be conducted to further explore these effects.

Inclusion of systematic maintenance procedures. As has already been discussed, some of the decline in performance during the follow up phase may have been due to the absence of a systematic maintenance program. Future studies should include systematic maintenance procedures, in line with what is typical clinical practice within an IBI centre.

Generalizability of results. Future studies should be conducted to explore the field effectiveness of stimulus equivalence in teaching additional skills other than

reading. Laboratory studies have been conducted on a variety of skills including math, geography, and money (LeBlanc et al., 2003; Maydak et al., 1995; Trace et al., 1977), and the applied literature should follow suit. If the results of the present study are generalizable to other skills as well then the implications for IBI programs could be substantial.

Conclusions

In summary, the present study has provided preliminary evidence for the field effectiveness of stimulus equivalence to teach reading skills to children with Autism. This study has highlighted the potential efficiency and economy of this approach. Participants came into the study able to demonstrate two different tasks for nine words. They were taught one additional task for each of the nine words. Up to eight new skills emerged for each of the nine words, without any additional teaching. At the completion of this study therefore, an equivalence network of up to 99 relations was present for each participant. Up to 72 of these relations emerged in the absence of any teaching. The results of this study are quite promising and suggest merit for this line of research in applied settings.

References

- Aaron, P.G., Frantz, S.S., & Manges, A.R. (1990). Dissociation between comprehension and pronunciation in dyslexic and hyperlexic children. *Reading and Writing, 2*(3), 243-264.
- Allen, D. A. (1988). Autistic Spectrum Disorders: Clinical presentation in preschool children. *Journal of Child Neurology, 3*, 48-56.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association.
- Bowder, D.M. & Lalli, J.S. (1991). Review of research on sight-reading instruction. *Research in Developmental Disabilities, 12*, 203-228.
- Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., & Hill, B.K. (1996). *Scales of Independent Behavior-Revised, Short Form*. Rolling Meadows, IL: Riverside Publishing.
- Catts, H.W., & Kamhi, A.G. (2005). *Connections between language and reading disorders*. Lawrence Erlbaum Associates.
- Cautilli, J.D., Hancock, M.A., Henry, C.A., & Tillman, C. (2002). Behavior Therapy and autism: Issues in diagnosis and treatment. *Behavior Analyst Today, 3*, 230-242.
- Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied Behavior Analysis*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall
- Cowley, B.J., Green, G., & Braunling-McMorrow, D. (1992). Using stimulus equivalence procedures to teach name-face matching to adults with brain injuries. *Journal of Applied Behavior Analysis, 25* (2), 461-475.
- Cuvo, A.J., Veitch, V.D., Trace, M.W., & Konke, J.L. (1978). Teaching change computation to the mentally retarded. *Behavior Modification, 2*, 531-548.

- DeLeon, I. & Iwata, B. (1996). Evaluation of a multiple stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, 29, 519-533.
- DeRose, J.C., DeSouza, D.G., & Hanna, E.S. (1996). Teaching reading and spelling: exclusion and stimulus equivalence. *Journal Of Applied Behavior Analysis*, 29, 451-469.
- Devany, J.M., Hayes, S.C., & Nelson, R.O. (1986). Equivalence class formation in language-able and language disabled children. *Journal of the Experimental Analysis of Behavior*, 46, 243-257.
- DiLalla, D.L. & Rogers, S.J. (1994). Domains of the Childhood Autism Rating Scale: Relevance for diagnosis and treatment. *Journal of Autism and Developmental Disorders*, 24, 115-128.
- Driscoll, C., & Kemp, C. (1996). Establishing the equivalence of single word reading and language in children with disabilities. *Journal of Intellectual and Developmental Disability*, 21, 115-139.
- Eaves, R. C., & Milner, B. (1993). The criterion-related validity of the Childhood Autism Rating Scale and the Autism Behavior Checklist. *Journal of Abnormal Child Psychology*, 21, 481-491.
- Eikeseith, S. & Jahr, E. (2001). The UCLA reading and writing program: a evaluation of the beginning stages. *Research in Developmental Disabilities*, 22, 289-307.
- Eikeseth, S. & Smith, T. (1992). The development of functional and equivalence classes in high functioning autistic children: the role of naming. *Journal of the Experimental Analysis of Behavior*, 58, 123-133.
- Elder, J.H. (2002). Current treatments in Autism: examining scientific evidence and clinical implications. *Journal of Neuroscience Nursing*, 34(2), 67-73

- Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in Autistic and Dyslexic children. *British Journal of Developmental Psychology*, 1, 329-342.
- Garfin, D. G., McCallon, D., & Cox, R. (1988). Validity and reliability of the Childhood Autism Rating Scale with autistic adolescents. *Journal of Autism and Developmental Disorders*, 18, 367-378.
- Green, G. (1990). Differences in development of visual and auditory-visual equivalence relations. *American Journal on Mental Retardation*, 95 (3), 260-270.
- Green, G. (2001). Behavior analytic instruction for learners with Autism: Advances in stimulus control technology. *Focus on Autism and Other Developmental Disabilities*, 16, 72-85.
- Green, G. & Saunders, R.R. (1998). Stimulus Equivalence. In Lattal, K.A. & Perone, M. (Eds.), *Handbook of Research Methods in Human Operant Behavior*. New York, NY: Plenum Press.
- Hall, S.S., DeBernardis, G.M., & Reiss, A.L. (2006). The acquisition of stimulus equivalence in individuals with Fragile X syndrome. *Journal of Intellectual Disability Research*, 50, 643-651.
- Horner, R.D. & Baer, D.M. (1978). Multiple-probe technique: a variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11, 189-196.
- Horner, R.H., Day, H.M., Sprague, J.R., O'Brien, M., & Heathfield, L.T. (1991). Interspersed requests: a nonaversive procedure for reducing aggression and self-injury during instruction. *Journal of Applied Behavior Analysis*, 24, 265-278.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.
- Lazar, R.M., Davis-Lang, D., & Sanchez, L. (1984). The formation of visual stimulus equivalence in children. *Journal of The experimental Analysis of Behavior*, 41, 251-266.

- LeBlanc, L.A., Miguel, C.F., Cummings, A.R., Goldsmith, T.R., & Carr, J.E. (2003). The effects of three stimulus-equivalence testing conditions on emergent US geography relations of children diagnosed with Autism. *Behavioral Interventions*, 18, 279-289.
- Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, 55(1), 3-9.
- Lynch, D.C., & Cuvo, A.J. (1995). Stimulus equivalence instruction of fraction-decimal relations. *Journal of Applied Behavior Analysis*, 28, 115-126.
- Mackay, H.A. (1985). Stimulus equivalence in rudimentary reading and spelling. *Analysis and Intervention in Developmental Disabilities*, 5, 373-387.
- Maurice, C., Green, G., & Luce, S. (Eds.). (2002). *Behavioral intervention for young children with autism: A manual for parents and professionals* (pp. 29-44). Austin, TX: Pro-Ed.
- Maydak, M., Stromer, R., Mackay, H.A., & Stoddard, L.T. (1995). Stimulus classes in Matching-To-Sample and sequence production: The emergence of numeric relations. *Research in Developmental Disabilities*, 16, 179-204.
- McEachin, J.J., Smith, T., & Lovaas, O.I. (1993). Long term outcome for children with Autism who received early intensive behavioural treatment. *American Journal on Mental Retardation*, 97, 359-372.
- Melchiori, L.E., DeSouza, D.G., & DeRose, J.C. (2000). Reading, equivalence, and recombination of units: A replication with students with different learning histories. *Journal Of Applied Behavior Analysis*, 33, 97-100.
- Nation, K., Clarke, P., Wright, B., & Williams, C. (2006). Patterns of reading ability in children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 36, 911-919.

- Nation, K., & Norbury, C.F. (2005). Why reading comprehension fails: insights from developmental disorders. *Top Language Disorders*, 25, 21-32.
- Newschaffer, C.J., Croen, L.A., Daniels, J., Giarelli, E., Grether, J.K., Levy, S.E., Mandell, D.S., Miller, L.A., Pinto-Martin, J., Reaven, J., Reynolds, A.M., Rice, C.E., Schendel, D., & Windham, G. (2007). The epidemiology of Autism spectrum disorder. *Annual Review Of Public Health*, 28, 235-258.
- Perry, A. & Freeman, N. (1996, March). Psychometric properties of the Childhood Autism Rating Scale. *Poster presented at the Ontario Association on Developmental Disabilities Research Special Interest Group Research Day, Toronto, ON.*
- Perry, A., Condillac, R.A., Freeman, N., Dunn-Geier, J., & Belair, J. (2005). Multi-site study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders*, 35 (5), 625-634.
- Perry, A., Cummings, A., Geier, J.D., Freeman, N.L., Hughes, S., LaRose, L., Managhan, T., Reitzel, J., & Williams, J. (2008). Effectiveness of intensive behavioural intervention in a large, community-based program. *Research in Autism Spectrum Disorder*, 2, 621-642.
- Rinaldi, L., Sells, D., & McLaughlin, T.F. (1997). The effects of reading racetracks on the sight word acquisition and fluency of elementary students. *Journal of Behavioral Education*, 7, 219-233.
- Saunders, R.R. & Green, G. (1999). A discrimination analysis of training-structure effects on stimulus equivalence outcomes. *Journal of The Experimental Analysis of Behavior*, 72, 117-137.

- Saunders, R.R., Wachter, J., Spradlin, J.E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of The Experimental Analysis of Behavior*, 49, 95-115.
- Schopler, E. C., Reichler, R., & Renner, B. (1988). The Childhood Autism Rating Scale (CARS). Los Angeles, CA: Western Psychological Services.
- Schreibman, L. (2000). Intensive behavioral/psychoeducational treatments for Autism: Research needs and future directions. *Journal of Autism and Developmental Disorders*, 30, 373-378.
- Seltzer, M.M., Shattuck, P., Abbeduto, L., Greenberg, J.S. (2004). Trajectory of development in adolescents and adults with Autism. *Mental Retardation and Developmental Disabilities Research Reviews*, 10, 219-236.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14, 5-13.
- Sidman, M. (1994). Equivalence Relations and Behavior: A research story. Boston, MA: Authors Cooperative Inc.
- Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 74, 127-146.
- Sidman, M. (2009). Equivalence relations and behavior: an introductory tutorial. *The Analysis of Verbal Behavior*, 25, 5-17.
- Sidman, M., & Cresson, O. (1973). Reading and crossmodal transfer of stimulus equivalence in severe retardation. *American Journal of Mental Deficiency*, 77, 515-523.
- Sidman, M., Cresson, O., & Willson-Morris, M. (1974). Acquisition of Matching-To-Sample via mediated transfer. *Journal of The Experimental Analysis of Behavior*, 22, 261-273.

- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. Matching-To-Sample: an expansion of the testing paradigm. *Journal of The Experimental Analysis Of Behavior*, 37, 5-22.
- Simpson, R.L. (2005). Evidence-based practices and students with Autism Spectrum Disorder. *Focus on Autism and Other Developmental Disabilities*, 20, 140-149.
- Smith, T. (2005). The appeal of unvalidated treatments. In J.W. Jacobson, R.M. Foxx and J.A. Mulick (Eds.), *Controversial Therapies for Developmental Disabilities: Fad, Fashion and Science in Professional Practice*. (pp.19-30). Mahwah, NJ: Lawrence Erlbaum Associates
- Stokes, T. & Baer, D. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10(2), 349-367.
- Stromer, R., & Mackay, H.A. (1992). Spelling and emergent picture-printed word relations established with delayed identity matching to complex samples. *Journal of Applied Behavior Analysis*, 25, 893-904.
- Stromer, R., Mackay, H.A., Stoddard, L.T. (1992). Classroom applications of stimulus equivalence technology. *Journal of Behavioral Education*, 2, 225-256.
- Tachimori, H., Osada, H., & Kurita, H. (2003). Childhood Autism Rating Scale, Tokyo version for screening pervasive developmental disorders. *Psychiatric and Clinical Neurosciences*, 57, 113-118.
- Trace, M.W., Cuvo, A.J., & Criswell, J.L. (1977). Teaching coin equivalence to the mentally retarded. *Journal of Applied Behavior Analysis*, 10, 85-92.
- Vause, T., Martin, G.L., Marion, C. & Sakko, G. (2005). Teaching equivalence relations to individuals with minimal verbal repertoires: Are visual and auditory discriminations predictive of stimulus equivalence? *The Psychological Report*, 55, 197-218.
- Wetherby, B., Karlan, G.R., & Spradlin, J.E. (1983). The development of derived

stimulus relations through training in arbitrary-matching sequences. *Journal of The Experimental Analysis of Behavior*, 40, 69-78.

Wing, L. (1988). The continuum of autistic disorders. In E. Schopler & G. M. Mesibov (Eds.), *Diagnosis and assessment in autism* (pp. 91-110). New York: Plenum Press.

Child's Name:

Date/Initials: _____ **Phase:** _____ (baseline or post test)

Please run tests in the order of the columns. Run each column for a number of words (3-4) before moving on to the next column.

For the target word set, for columns 1-7 complete a minimum of two trials per testing session and a minimum of three sessions. For columns 8 and 9 complete a minimum of three trials per session and a minimum of three sessions.

For the word sets which is not currently the target set complete a single probe trial (one session only)

[illegible]

2. Data sheets for modified post tests

Child's Initials: _____

Modified Data Sheet for Baseline and Post Test PART 1

Record + for correct responses and – for incorrect responses.

Please run tests in the order of the columns. Run each column for all the words in the set (3) before moving on to the next column.

Complete two trials per testing session and across three sessions.

Date/Initials: _____ Original Data collector or
IOA _____

TARGET	Match PICTURE to dictated name (Touch_____)	Match WRITTEN WORD to dictated name (Touch_____)	Match written word to picture	Match picture to written word
1				
2				
3				

Modified Data Sheet for Baseline and Post Test PART 2

Record + for correct responses and – for incorrect responses.

Please run tests in the order of the columns. Run each column for all the words in the set (3) before moving on to the next column.

Complete two trials per testing session and across three sessions.

Date/Initials: _____

Original Data collector or IOA _____

Child's Initials: _____

TARGET	Oral naming of PICTURE (What's this?)	Match picture to written word	Oral naming of WRITTEN WORD
1			
2			
3			

Modified Data Sheet for Baseline and Post Test PART 3

Record + for correct responses and – for incorrect responses.

Please run tests in the order of the columns. Run each column for all the words in the set (3) before moving on to the next column.

Complete three trials per testing session and across three sessions

Date/Initials: _____

Original Data collector or IOA _____

Child's Initials: _____

TARGET	Match Picture to Dictated name (Touch ____)			Match Dictated name to Picture (language master)		
1						
2						
3						

Modified Data Sheet for Baseline and Post Test PART 4

Record + for correct responses and – for incorrect responses.

Please run tests in the order of the columns. Run each column for all the words in the set (3) before moving on to the next column.

Complete three trials per testing session and across three sessions

Date/Initials: _____

Original Data collector or IOA _____

TARGET	Match Written word to Dictated name (Touch ____)			Match Dictated name to written word (language master)		
1						
2						
3						

Modified Data Sheet for Baseline and Post Test PART 5

Record + for correct responses and – for incorrect responses.

Please run tests in the order of the columns. Run each column for all the words in the set (3) before moving on to the next column.

Complete two trials per testing session and across three sessions

Date/Initials: _____

Original Data collector or IOA _____

TARGET	Match Name to Name (vocal imitation)	Match picture to picture	Match word to word
1			
2			
3			

3. Data Sheet for teaching

Child's Name: _____ Date/Initials: _____

Program Sight Reading Program Child will match written words to dictated name in a messy array of 6 TARGET: _____	Start S ^P T % _____	Start R ⁺ FR VR _____	Instructional Trials and Staff Initials										End S ^P T % _____	End R ⁺ FR VR _____
	Prompt type (Phy / Ver / Vis / Ges / Pos / W-In / T Delay)		1	2	3	4	5	6	7	8	9	10	Notes	
	Prompt amount (%)		100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5	100 75 50 25 5		
Stimuli 1 target word, 5 distracters of similar length and similar letters	Learner's response		X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	X PC ✓	S ^{R1} Social Praise S ^{R2} _____ S ^{R3} _____	

4. Modified baseline data sheet and data sheet for follow up

Child's Name: _____ **Date/Initials:** _____ **Original Data collector or IOA** _____

[illegible]

5. Generalization data sheet

Generalization Data Sheet:**Date:**_____**Child's Name:**_____**Initials:**_____

Instructions: Complete a single probe for each word for each of the steps below using novel stimuli not presented during any of the previous teaching or testing phases. (Stimuli have been provided)

Target Word	Step 4 Oral Naming of Written word (Label)		Step 6 Match picture to written word		Step 7 Match written word to picture		Step 9 Match written word to dictated name (Touch)	
Word Set 1								
1								
2								
3								
% out of 6								
Word Set 2								
1								
2								
3								
% out of 6								
Word Set 3								
1								
2								
3								
% out of 6								

Appendix B: Word Selection results
Table B2:

Word Selection Data for Participant 1

Target Word	Oral Naming of Written word	Oral Naming of Picture	Match picture to written word	Match written word to picture	Matching picture to dictated name	Match written word to dictated name	Word chosen
1 TWO	---	+++	---	---	+++	---	YES
2 BEE	-	+	+				
3 DOG	-----	+++	--++--	-----+	+++	-----	
4 POP	---	+	+---	---	Error made	---	
5 BUS	----	+	+---	---	Error made	---	
6 ONE	-----	+	+-----+	---	Error made	+--	YES
7 EAR	----	+	----	++--			
8 CAT	---	+++	---	---	+++	---	
9 OWL	----	+	++--	-+-			
10 PIG	-----	++	+-----	-+---+-	Error made	-	
11 CUP	-----	++	-+-----+	-			YES
12 BED	---	+++	---	---	+++	---	
13 SIX	---	+++	--++	--	++	--	
14 DUCK	-----	+++	--+-+	---	++	---	
15 BEAR	---	Error made	---	-+-	Error made	+--	
16 LION	---	Error made	---	---	Error made	---	YES
17 FROG	---	Error made	---	---	Error made	+--	
18 BIKE	-----	+++	+-----	-----	+++	-----	
19 CAKE	-----	+++	---	++			
20 BOOK	-----	+++	+-----	-----	+++	-----	
21 BABY	---	+++	---	---	+++	---	YES
22 FORK	-----	-++++ ++	-----	-----	++++++ +	+-----	YES
23 FIVE	-----	+++	-----	-----	+++	-+-----	YES
24 WOLF	---	+++	---	---	+++	---	YES

Table B4:
Word Selection Data for Participant 2

Target Word	Oral Naming of Written word	Oral Naming of Picture	Match picture to written word	Match written word to picture	Matching picture to dictated name	Match written word to dictated name	Word chosen ?
1 BED	---	+++	---	- +			
2 ANT	---	+++	---	---	+++	---	YES
3 CUP	---	+++	---	---	+++	---	YES
4 DOG	---	+++	---	---	+++	---	YES
5 EAR	---	+++	---	---	+++	---	YES
6 OWL	---	+++	---	-- +			
7 PEN	---	---					
8 TWO	---	+++	- +				
9 BABY	---	+++	+				
10 CAKE	---	+++	- +				
11 DRUM	---	- + +	---	---	+++	---	
12 FISH	---	+++	---	---	+++	---	YES
13 LION	---	+++	---	---	+++	---	YES
14 MILK	---	+++	---	---	+++	---	YES
15 STAR	---	+++	---	---	+++	---	YES
16 DOOR	---	+++	---	---	+++	---	YES

Table B6:

Word Selection Data for Participant 3

Target Word	Oral Naming of Written word	Oral Naming of Picture	Match picture to written word	Match written word to picture	Matching picture to dictated name	Match written word to dictated name	Word chosen ?
1 BED	---	+++	---	---	+++	--+	YES
2 CAT	---	+++	---	---	+++	---	
3 BUS	---	+++	---	---	+++	---	
4 DOG	---	+++	-+				YES
5 EAR	---	+++	---	+			
6 EGG	---	+++	-+				
7 PIG	---	+++	---	---	+++	---	YES
8 BABY	---	+++	-+				YES
9 BEAR	---	+++	---	---	+++	---	
10 CAKE	---	+++	-+				
11 DRUM	---	+++	+				YES
12 BOOK	---	+++	---	---	+++	---	
13 FORK	---	+++	---	---	+++	---	
14 BOAT	---	+++	---	---	+++	---	YES
15 CORN	---	+++	-+				YES
16 DUCK	---	+++	---	---	+++	---	
17 BIKE	---	+++	---	---	+++	---	
18 GLUE	---	+++	---	---	+++	---	YES

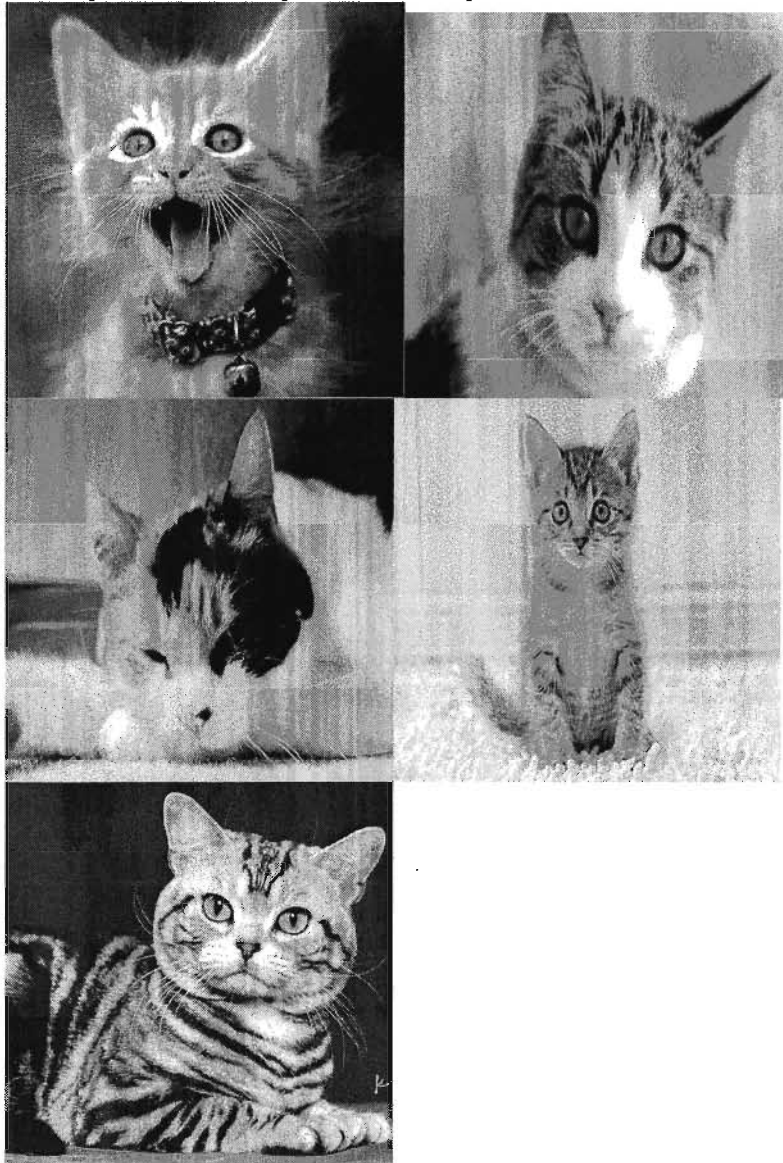
Table B8:

Participant 4 Word Selection

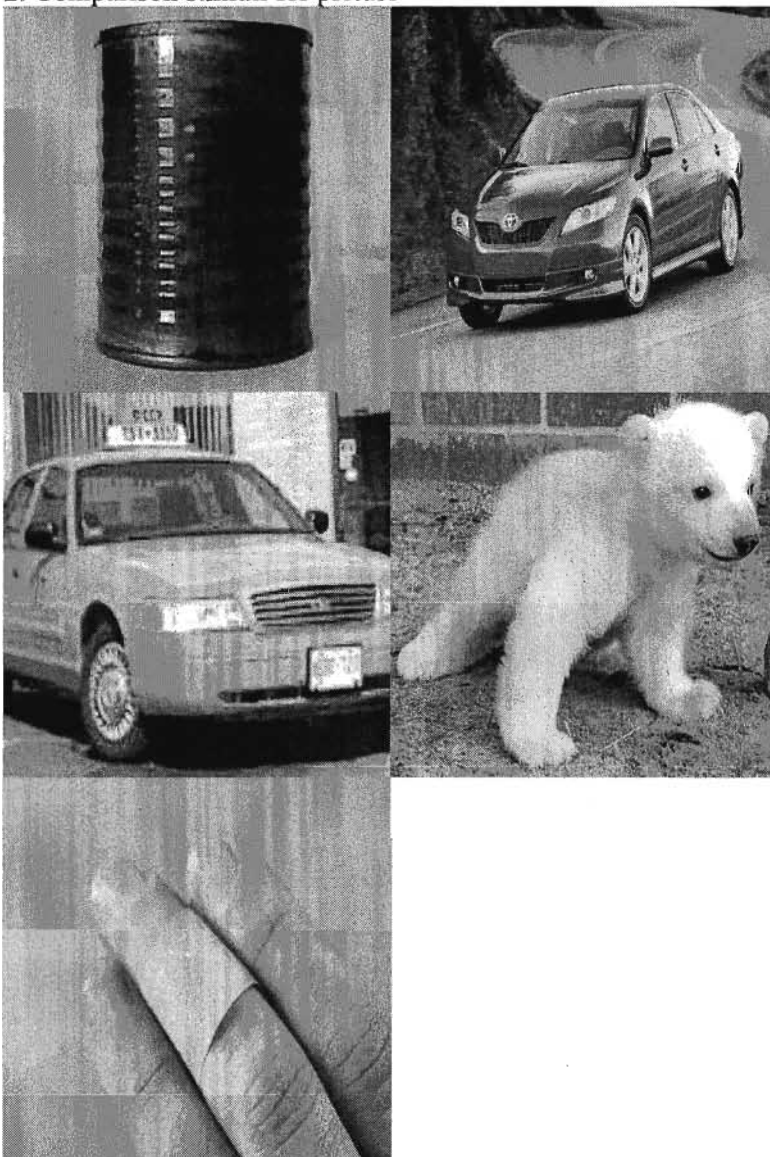
Target Word	Oral Naming of Written word	Oral Naming of Picture	Match picture to written word	Match written word to picture	Matching picture to dictated name	Match written word to dictated name	Word chosen ?
1. Dog	---	+++	---	---	-		
2. Two	-----	+++ +++	-----	-----	-+++	---	YES
3. Bird	---	+++	---	-+			
4. Door	---	-					
5. Milk	---	-					
6. Fork	---	-					
7. Tree	---	+++	---	---	+++	---	YES
8. Star	-----	+++++	-----	+ - +			
9. Bowl	---	++-					
10. Bike	---	+++	---	---	+++	---	YES
11. Lion	---	+++	---	---	+++	---	YES
12. Bed	---	-					
13. Bag	---	+++	---	---	+-		
14. Ear	---	-					
15. Oval	---	++-					
16. Cup	---	-					
17. Fish	---	+++	---	---	+-		
18. Cake	---	+++	---	-+			
19. Cat	---	+++	---	---	+++	---	YES
20. Boat	---	+++	---	---	+++	---	YES
21. Frog	---	+++	--+				
22. Four	---	+++	+				
23. Six	---	+++	---	+			
24. Bus	---	-					
25. Baby	---	+++	---	---	+++	---	YES
26. Five	---	+++	---	---	+++	---	YES
27. Book	---	+++	---	---	+++	---	YES

Appendix C: Example of stimuli used:

1. Sample stimulus for picture: 5 exemplars



2. Comparison stimuli for picture



3. Sample stimulus for written word-5 exemplars

Cat

CAT

cat

Cat

Cat

4. Comparison stimuli for written word (sample for each font was used, only one is provided here as an example).

Can

Car

Cab

Cut

Cub